THE ECONOMIC JOURNAL

The Economic Journal, **133** (*February*), 888–903 https://doi.org/10.1093/ej/ueac078
Advance Access Publication Date: 12 November 2022

THE LONG-TERM EFFECTS OF STUDENT ABSENCE: EVIDENCE FROM SWEDEN*

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

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 outcomes. We address this question by combining historical and administrative records for cohorts of Swedish individuals born in the 1930s. We find that elementary school absence significantly reduces contemporaneous academic performance, final educational attainment and labour 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 labour market performance.

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. 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 little causal evidence of the effect of absence on socio-economic outcomes. The few papers that credibly establish such evidence find that absences in elementary and secondary school have a detrimental impact on academic achievement and school graduation (Goodman, 2014; Aucejo and Romano, 2016; Liu *et al.*, 2021). While these papers are important in going beyond the correlational evidence, they all focus on educational outcomes in the United States. Much remains to be known as to whether (*i*) impacts on educational outcomes reflect human capital losses that translate into long-term outcomes and (*ii*) whether these impacts generalise to other contexts.

* Corresponding author: Sarah Cattan, Institute for Fiscal Studies, 7 Ridgmount Street, WC1E 7AE London, UK. Email: sarah_c@ifs.org.uk

This paper was received on 1 March 2022 and accepted on 17 October 2022. The Editor was Barbara Petrongolo.

The authors were granted an exemption to publish their data because access to the data is restricted. However, the authors provided the Journal with temporary access to the data, which allowed the Journal to run their codes. The codes are available on the Journal repository. The data and codes were checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: https://doi.org/10.5281/zenodo.7064130.

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), Lund University, University of Essex as well as participants of EEA 2015, ESPE 2015, IWAEE 2015, VfS 2015 and the Essen Health Conference 2015. For collecting and digitising the data used here we are indebted to our colleagues in Essen and Lund as well as a vast team of research assistants. Earlier versions of this paper circulate as IZA Discussion Paper 10995 (September 2017), ISF Working Paper 06/21 (February 2021) and DICE Discussion Paper 383 (March 2022). Sarah Cattan gratefully acknowledges financial assistance from the British Academy Postdoctoral Fellowship pf140104 and the ESRC-funded Centre for the Microeconomic Analysis of Public Policy (ES/T014334/1). Therese Nilsson gratefully acknowledges financial support from the Swedish Research Council (dnr 2019-03553), the Gyllenstierna Krapperup Foundation, KEFU and the Crafoord foundation (dnr 20190685). 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.

This paper provides evidence of the impact of student absence on educational achievement, labour market outcomes and mortality in the context of Sweden. We construct a unique panel following a representative sample of cohorts born 1930–5, which links digitised records of absence and educational performance in elementary school with census and tax register data on socio-economic outcomes throughout the life cycle. To our knowledge, this is the first dataset that allows for an analysis of the long-term impact of absences. This is an important innovation in light of evidence 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 rich data, analysing the short- and long-term impacts of student absences requires a robust strategy to tease out their causal impacts from the vast array of unobserved confounding factors. Indeed, students who miss school may have poorer health, less engaged parents and/or less inspiring teachers, which could all lead to spurious correlations between absence, education and labour market outcomes. To deal with the endogeneity of absence, we exploit two features of our data. For the short-term analysis of student absences on school performance, we exploit within-student, between-grade variation in absence and performance at two time points (grade 1 and grade 4). For the long-term analysis, we exploit the presence of siblings in our dataset, and use within-family variation in absence and long-term outcomes to purge the correlation between them from all family-level time-invariant factors. We present a series of robustness tests supporting the identification strategy and estimate bounds following Oster (2019), suggesting that biases associated with our estimates are unlikely to be either statistically or economically large.

Our analysis yields two main findings. First, we find a negative and significant impact of student absence on academic performance in elementary school equivalent to 4.5% of a standard deviation (SD) for ten days of absence. This impact is similar to estimates presented for more recent cohorts of elementary school children growing up in the United States (Goodman, 2014; Aucejo and Romano, 2016). Linking school performance to later-life earnings, we also express the effect of student absence in terms of the long-term earnings potential.

Second, estimates of the long-term effects of absence are most pronounced for our income measures: income measured at ages 35–40 and pension income, which is a proxy for lifetime earnings. Ten annual days of absence in elementary school are estimated to decrease these income measures by 1%–2%. The estimates of the impacts 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 complete secondary education. The results are consistent with a model of skill accumulation, where early losses in human capital grow into 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.

Our paper relates to a broad literature examining the impact of instructional time on educational achievement and adult outcomes. Although school absence is an important determinant of the total amount of time spent in school, most studies exploit exogenous variation in the length of the school year as a source of exogenous variation in instructional time (see, e.g., Pischke, 2007; Leuven *et al.*, 2010; Agüero and Beleche, 2013; Fischer *et al.*, 2019 using laws and law changes; Fitzpatrick *et al.*, 2011; Carlsson *et al.*, 2015 using variation in test dates; Marcotte and Hemelt, 2008; Marcotte and Hansen, 2010; Hansen, 2011 using unscheduled school closures).

Regarding individual absence from school, three studies, Goodman (2014), Aucejo and Romano (2016) and Liu *et al.* (2021), analyse the effects of school absence on educational achievement in the United States. Using Massachusetts data (2003–10) for students attending grade 3 onwards and North Carolina data (2006–10) for grade 3 to 5 students, respectively, Goodman (2014) and Aucejo and Romano (2016) showed that school results are negatively affected by absence. Both studies control for institutional heterogeneity using school, teacher and individual fixed effects, as we do. Liu *et al.* (2021) estimated the impact of absences during secondary school on educational outcomes using data from a Californian school district (2002–13). They used between-subject, within-individual variation in absence to identify the impact of absence on contemporaneous achievement, and estimate bounds around the impact of absence on high school graduation and college enrolment using the method of Oster (2019).

We contribute to the above literature 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 outside the United States, and on adult and later life outcomes. The literature examining the effect of region- or school-level changes in instructional time suggests that institutions and the educational system are relevant factors for observed effects (see, e.g., Pischke, 2007; Gathmann *et al.*, 2015; Galama *et al.*, 2018). Sweden makes a particularly interesting case in comparison to the United States, given that its labour market was characterised by active labour market policies and compressed wages (Erixon, 2008), and embedded in a Social Democratic welfare state providing comprehensive social insurance against health and social risks that workers face (Bergh, 2014).

Our results show that these innovations to the literature matter for our understanding of the impact of school absences. Considering effects throughout the life cycle sheds new light on previous findings regarding the role of school absence, and indicate that the short-term human capital losses manifested as impacts on test scores translate into long-term penalties in the labour market.

1. Background and Data

1.1. *The Swedish Education System*

In the 1930s and 1940s, all children in Sweden were required to attend public school, *Folkskola*, starting at the age of seven. Education was free and co-educational. The first four grades of *Folkskola*, which we refer to as elementary school, were mandatory. Admissions to secondary school were very selective and depended on academic performance, and the minority of students who progressed to secondary school generally matriculated after grade 4. Students not pursuing post-compulsory schooling had to remain in *Folkskola* for another two years.[1]

The responsibility of providing elementary education was decentralised to the 2,400 school districts, but the Ministry of Ecclesiastical Affairs provided nationwide standards, for example for curriculum design and grading. Three academic subjects were taught in elementary school: *math*, *reading and speaking*, and *writing*. The government established grading principles, dictating that teachers should reward the quality of knowledge and regularly take notes to ensure that grades reflected performance through the year (SOU, 1942). Online Appendix A provides details on the school system.

[1] During elementary school, our cohorts were exposed to two educational reforms, rolled out across municipalities between 1936 and 1948. One reform expanded instructional time per school year from 34.5 to 36.4 and 39 weeks. We control for the number of days of the school year in all specifications. The second reform introduced a mandatory seventh grade of *Folkskola*. While about half of the individuals in our sample were affected by the reform, only 6% of the sibling pairs were affected differently. Controlling for the reform does not change our point estimates; see column (6) of Table 4.

As their main organisational tool, teachers kept daily records of students' performance, absence and reasons for absence in an exam catalogue (Online Appendix Figure A1). For every school year, teachers noted the total days of absence by type and final grades by subject. To conduct our analysis, we digitised end-of-year information from these catalogues and linked this information to administrative data on long-term outcomes.

1.2. *Data Sources*

The base data are a census of children born 1930–5 in a representative sample of 133 out of about 2,400 Swedish parishes; see Bhalotra *et al.* (2017). We construct our analysis datasets by combining this information with a number of historical and administrative sources described below. We provide information about the representativeness of the sample in Online Appendix B. Online Appendix Figure B1 shows the spatial distribution of our parishes, and Online Appendix Table B1 compares them to all parishes in Sweden in terms of 1930 population census information. Online Appendix B.2 defines all the variables included in the analysis.

1.2.1. *Church records*

The church records contain the child's name, gender, the date, parish and location (hospital or home) of birth, whether the birth was a singleton and the mother's marital status at the child's birth. These records also include information on the parents' names and occupations at the time of the child's birth. We create an indicator for maternal employment status and indicators for whether the father was an agricultural, a production or a service worker. We use all these variables as controls in our analysis (see summary statistics in Online Appendix Table B2).

1.2.2. *School records*

Schooling information comes from exam catalogues kept in historical archives, with yearly student-level absence by type (sickness and non-sickness) and grade points by subject (math, reading and speaking, writing) in grades 1 and 4 (years 1937–47). Grades 1 and 4 are pivotal, as grade 1 is the first occasion academic performance that can be observed, and grade 4 represents the last before some students proceed to secondary schooling. Grade points range from 1 (lowest) to 15 (highest); see Online Appendix Table A1. To facilitate interpretation, we standardise the grade points to have mean 0 and SD 1.

School records were matched to church records using information on parish, birth date and full name. We mitigate data lost on account of migration out of the birth parish by tracking migrants and collecting school records from destination parishes.[2]

1.2.3. *Final education*

Information on highest educational level completed comes from the 1970 population census (SCB, 1972), when individuals are aged 35–40. We create a binary indicator for whether an individual attains more than *Folkskola*.

1.2.4. *Labour market outcomes*

Information about employment at ages 25–30 and 35–40 comes from the 1960 and 1970 population censuses (SCB, 1962; 1972). For each age group, we construct employment variables that

[2] Online Appendix B.3 describes the matching and how we trace individuals who move to another school parish. Online Appendix Figure B2 gives the degree of selection bias possibly arising from attrition. Online Appendix Table B3 shows balancing results for the full sample and the exam catalogue sample in terms of socio-demographic characteristics. Online Appendix Table B4 indicates that linear regression estimates of absence on grade points do not differ substantially across samples.

Table 1. *Descriptive Statistics on Long-Term Outcomes.*

	(1)	(2)	(3) Mean	(4)	(5)	(6)
	Age range	All	Female	Male	# obs	% female
Education						
More than *Folkskola* (in %)		11.8	12.1	11.4	5,976	49.6
Employed (in %)						
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)						
Labour market income 1970	35–40	19,275	10,039	28,475	5,886	49.9
Pensions 2003–8, if > 0	68–78	160,237	138,568	183,468	4,772	51.7

Notes: The numbers refer to the individuals in the siblings sample (second row of Online Appendix Table B5) 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 1970 census, but likely reflects 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 and takes the value 1 if the individual is employed, and 0 otherwise. Labour market income 1970 is based on the 1970 census; unemployment enters the labour market income as zero. Pensions information is taken from tax registers and averaged over 2003–8. Zero pensions are dropped (very few cases). Labour market income and pensions are measured in Swedish krona (current values).

equal 1 if the individual works at least part time, and 0 otherwise. From the 1970 census we also have labour market income when individuals were in prime working age. The income measure has imputed zeros for individuals who were not employed.

1.2.5. *Pension income*

From tax registers we measure average annual pensions 2003–8, when our youngest cohort is aged 68–79. Because. for our cohorts, full pensions require 30 years of contributions and is based on the 15 highest income years (Sundén, 2006), this measure is a proxy for lifetime earnings. That pension income is less sensitive to year-to-year fluctuations in labour supply than annual earnings is a desirable feature, especially for women.[3]

1.2.6. *Death records*

Church records and the Swedish Death Index (Federation of Swedish Genealogical Societies, 2014) provides the exact date of death for all individuals that passed away. We generate indicators for whether the individual died before 1960, 1970 and 2003.

For adult outcomes, we match individuals in the matched schooling data to the different registers using either full name, sex, birth date and parish of birth, or a unique social security number (see Online Appendix B.4 for details about the matching procedure). Online Appendix Table B5 reports the numbers of individuals, siblings and families included in our analyses.

1.3. *Descriptive Statistics*

Table 1 presents descriptive statistics for our long-term outcomes, and Online Appendix C.1 presents descriptive statistics for absence by type and school achievements. In grade 1, the average number of missed days is 11.1 days (median seven days) versus 11.6 (median eight days)

[3] Online Appendix D provides details about the pension system and related rules.

in grade 4. Despite a very different context, 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 numbers of absence days (see Online Appendix Figure C1 for the overall variation, and Online Appendix Figure C2 for the within-family and within-individual variations), but unlike Goodman (2014), who excluded observations with more than 60 days of absence, we do not cap absence days.[4] 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 grade 4.

Average grade point in math, reading and writing, and speaking is higher in grade 4 than in grade 1. In line with national guidelines (cf. SOU, 1942), only 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 (Online Appendix Figures C3–C4). The selective nature of the education system is reflected by the fact that only 12% of our sample has more than *Folkskola* education.

The cohorts studied were born between 1930 and 1935, and entered the labour market around 1950.[5] Online Appendix C.2 provides insights into the most prevalent occupations of our cohorts, and shows that there was a high degree of gender segmentation in the labour market (see Online Appendix Table C2 and Online Appendix Figure C6). Employment and earnings measured at ages 25–30 and 35–40 reflect such segmentation. As expected, pensions are more equal between genders than earnings are.

Our main outcomes are all negatively correlated with total days of absence (see Online Appendix Figures C5 and C7). 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. To start exploring the extent of selection, we look at how total days of absence vary across groups defined by different observable characteristics (see Online Appendix Figure C9). Overall, these statistics show limited signs of selection based on these observables. Whether students positively or negatively select into absence is also unclear. As we describe in the next section, our identification strategies exploit within-individual or within-family variation in absence and outcomes. We show in Online Appendix Table C1 that there is sufficient variation of this kind.

2. Empirical Strategy

2.1. *Identification of Short-Term Effects*

While the main focus of this paper is on the long-term effects of school absence, a natural starting point is to estimate the contemporaneous effect of absence on school performance. This also allows us to compare effects with those found for the current US context.

Our aim is to identify the effect of the number of days of absence in one year on academic achievement measured at the end of that year. We exploit the fact that we observe performance

[4] Instead we winsorise our data at the top 98th percentile of the distribution of days of sickness and non-sickness absence.

[5] The cohorts under review were born during the Great Depression and were in elementary school during World War II. Although the depression was not as severe in Sweden as it was in many other countries, we re-estimate our model for long-term outcomes, controlling for local economic conditions (poverty rate, taxable income per capita and the annual change of the latter) in the year of birth; see column (7) of Table 4. These estimates are extremely similar to our main estimates and, if anything, more precise. WWII is unlikely to have disrupted either the education or the archiving of school records. Sweden was neutral in WWII and there was an oversupply of teachers (Paulsson, 1946). We have not found historical sources suggesting that WWII caused disruptions in education and the probability that we found exam catalogues in archives was the same for years during and not during the War.

and absence in two grades (1 and 4) in order to control for individual-specific time-invariant unobservable factors. We estimate the equation

$$Y_{istg} = \beta_0 + \tau D_{ig} + \boldsymbol{Q}_{stg}{}'\beta_2 + S_{sg} + T_{tg} + \delta_i + u_{istg}, \tag{1}$$

where Y_{istg} is grade-g performance of a student i attending school s taught by teacher t, D_{ig} is the number of days student i was absent from school in a grade, $\boldsymbol{Q}_{stg}$ is a set of school/grade-specific controls, S_{sg} is a grade-g school fixed effect (FE), T_{tg} is a grade-g teacher FE, and δ_i is an individual FE.[6]

As argued by Bond and Lang (2013), which measure of educational achievement is used can significantly alter the conclusion of the analysis. To address this issue and provide economically meaningful interpretations of our estimated effects of absence on achievement, we anchor the grade point scale to pensions, a policy-relevant outcome measured in a meaningful metric (Swedish krona, SEK). Online Appendix E provides details on the anchoring procedure and Appendix Table E1 reports estimates of the anchoring equations.

The main identifying assumption in model (1) is that unobserved grade- and child-specific factors affecting student i's achievement in grade g are uncorrelated with the child's absence, conditional on the observables and fixed effects included in the model. The assumption is the same as in Aucejo and Romano (2016), but stronger than the assumption required in the design of Liu *et al.* (2021), who exploited within-grade, between-subject variation in absences to control for time-varying individual-level unobserved shocks.[7] Given the main focus of the paper on long-term outcomes and the fact that this strategy has been used in several related papers, we refer the reader to Online Appendix F.3 for a discussion of threats to its validity and robustness checks.

2.2. *Identification of Long-Term Effects*

Our main analysis focuses on the effect of school absences in elementary school on adult outcomes. We define our treatment variable as the average number of days of absence in grades 1 and 4. Since the average days of absence in elementary school are fairly stable across all grades, we can think of the treatment variable as closely measuring the average yearly number of days of absence in each year throughout elementary school.[8]

Our research design exploits the fact that we can identify siblings in our data. We propose to identify the impact of school absence on long-term outcomes from within-family variation and estimate the equation

$$W_{if} = \beta_0 + \tau D_i + \boldsymbol{X}_i{}'\beta_1 + \omega_f + v_{if},$$

where W_{if} is an outcome of individual i of family f; $\boldsymbol{X}_i$ is a set of individual-specific time-invariant controls, including demographic characteristics, grade-1 school identifier, grade-1

[6] The specification assumes a homogeneous effect of absence in grades 1 and 4 in order to interpret τ as the marginal effect of a day of absence. This assumption cannot be tested in the context of the individual FE model, but it can be tested in the sibling FE model. We report results of such a test in the lower panel of Online Appendix Table F6, which shows that the effects of absence in grade 1 and in grade 4 are statistically indistinguishable.

[7] On the other hand, Liu *et al.* (2021) required an absence of spillover effects across subjects for identification.

[8] We have not collected information on grades 2 and 3 systematically, but some exam catalogues cover students in several grades. Average days of absence in grades 2 (10.7 days) and 3 (11.9) are close to the averages for grades 1 (11.1) and 4 (11.6). We also estimate the model allowing for a different effect of grade-1 and grade-4 absences on long-term outcomes. In the sibling FE model, estimates of these two effects are not statistically different from each other (Online Appendix Table G2).

teacher identifier, as well as information about class size, lowest and highest grade taught to students in the same classroom and length of the school year in weeks, averaged across grades 1 and 4. We also control for an indicator for whether grade-1 teacher is the same as grade-4 teacher and an indicator for whether grade 1 school is the same as grade 4 school. Here ω_f is a family FE.[9]

The advantage of this strategy is that, in contrast with the bounding method used in Liu *et al.* (2021), it allows us to point identify the effect of absence on long-term outcomes. However, for the parameter τ to have a causal interpretation, it requires that any child-specific unobservable that does not have the same additive effect on outcomes for both siblings is uncorrelated with the child's absence. We discuss possible threats to identification in Subsection 3.3, where we present the results of a number of exercises to gauge the research design's validity.

3. Results

3.1. *Short-Term Effects of School Absence*

Column (1) of Table 2 reports ordinary least square (OLS) estimates of parameter τ in (1), and shows that one additional day of absence is significantly associated with a 0.51% of an SD decrease in performance. The individual FE estimate in column (2) is slightly smaller (0.0045), but also statistically significant at the 1% level.[10]

Assuming linearity, the effect of ten days of absence—the average in our sample—corresponds to 4.5% of an SD decrease in performance. Despite analysing absence in a very different context and literally in another century, our results measured in SD units are comparable to those in the related literature. Using recent US data, Goodman (2014) found an effect of 0.8% of an SD in math and English, and Aucejo and Romano (2016) found effects of 0.55% of an SD in math and 0.29% in reading. Online Appendix Figures F1 and F2 present evidence strongly suggesting that the relationship between absence and academic performance is linear. Moreover, we show that there are no significant differences in the short-term impact of absences between gender and across individuals from different socio-economic groups (Online Appendix F.2).

By anchoring student performance to pension incomes, we translate the short-term effect of absence on school performance into its effect on earnings potential. Table 2 (second and third rows) reports the results. In the individual FE specification, the impact of ten additional absence days on school performance translates into a non-significant decrease in earnings potential of SEK 16, equivalent to a 0.1% decrease in average pension income. We obtain similarly sized effects when anchoring by grade and gender (third row), and when using earnings at age 35–40, as the anchor.

Online Appendix F.3 presents results of tests we conduct to probe the validity of the individual FE strategy to recover the causal effect of absence on performance. These results strongly suggest that our short-term estimates of absence are likely robust to the presence of individual-level time-varying unobserved heterogeneity.

[9] Online Appendix Table C1 shows that a large fraction of variation in absence and performance in the data comes from within-family variation, which provides confidence that there is sufficient variation to consider using a sibling FE strategy.

[10] Online Appendix Table F1 reports estimates for control variables. Online Appendix Table F2 states negative and significant estimates of the effect of absences on performance for each of the three subjects. Online Appendix Table F3 shows that the estimates are unchanged when grade points are standardised before the aggregation. Online Appendix Table F4 gives the estimates when averaging over the math grade points and reading and speaking grade points (without writing). Online Appendix Table F5 gives the estimates when grades are coded on a seven-point scale.

Table 2. *Estimates of the Short-Term Effect of School Absence on Academic Performance.*

	(1) OLS	(2) Individual fixed effects
Panel A: average grade points in units of SD (mean 0, SD 1)		
Days of absence	−0.0051***	−0.0045***
	(0.0007)	(0.0013)
Panel B: average grade points in units of pension (anchoring by grade, mean pension in sample: 160,237 Swedish krona)		
Days of absence	−54.7344***	−16.1192
	(7.4068)	(20.0252)
Panel C: average grade points in units of pension, by gender (anchoring by grade and gender, mean pension in sample: 138,568 Swedish krona for women and 183,468 Swedish krona for men)		
Days of absence	−69.2581***	−37.5660**
	(9.2558)	(18.4195)
# observations	14,066	8,934
# individuals/families		4,467

Notes: Each cell reports the coefficient associated with days of absence in separate regressions where the dependent variable is the variable indicated in the first row of each panel. The dependent variable in panel A is average performance over math, reading and speaking, and writing, standardised to have mean 0 and SD 1. In panel B, it is a measure of average grade points in units of pensions, where the relationship between grade points and pensions is estimated separately for grade 1 and grade 4; see Online Appendix E. The third panel repeats the anchoring of grade points in units of pensions, but estimates the relationship between grade points and pensions separately by school grade and gender. Both OLS and individual FE models control for grade, range of grades instructed in the same classroom and length of the school year in weeks. The OLS specification also controls for gender, born out of wedlock, twin birth, mother employed at the time of birth, born in hospital, full sets of fixed effects for the year and month of birth, year and month interactions, age, parent's year of birth and family's socio-economic status based on the first digit of the Historical International Standard Classification of Occupations (HISCO) code (see van Leeuwen *et al.*, 2002) of the father. Standard errors clustered at the parish level are given in parentheses. Significance: $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

3.2. *Long-Term Effects of School Absence*

Table 3 reports the effects on long-term outcomes of the average days of absence across grades 1 and 4 in columns (1) (OLS) and (3) (sibling FEs). To ease interpretation, we also report the effect of ten days of absence relative to the mean of the outcome considered in columns (2) and (4). We calculate this by multiplying the coefficient by five (the treatment variable measures the average number of days of absence across two grades) and dividing it by the outcome mean.

A first observation is that the OLS and sibling FE estimates are similar both in terms of magnitude and statistical significance. This aligns with one of the conclusions from the short-term analysis that selection into absence based on unobservables may not be particularly prominent in our context.

Looking at the sibling FE estimates of Table 3, we find that ten days of absence in elementary school leads to a statistically insignificant 2.2% reduction in secondary school completion relative to baseline. Given the gender differences in educational attainment that existed during this period, impacts could vary by gender. Indeed, the effect of absence on secondary school completion is negative and statistically significant for men, while it is close to zero for women (Online Appendix Table G1). The negative effect of elementary school absence on secondary schooling for men is consistent with secondary school admissions depending on elementary school performance.

Table 3. *Long-Term Effects of School Absence.*

	(1) OLS Coefficient	(2) Rel. size	(3) Sibling fixed effects Coefficient	(4) Rel. size
More than Folkskola *(1 = yes)*				
Absence (average, grades 1 and 4)	−0.0007 (0.0005)	−0.0297	−0.0005 (0.0009)	−0.0215
Employment 1960 (1 = yes)				
Absence (average, grades 1 and 4)	−0.0012* (0.0007)	−0.0088	−0.0007 (0.0013)	−0.0053
Employment 1970 (1 = yes)				
Absence (average, grades 1 and 4)	−0.0008 (0.0007)	−0.0050	−0.0020 (0.0014)	−0.0133
Labour market income 1970				
Absence (average, grades 1 and 4)	−54.2973** (22.7088)	−0.0141	−80.7042** (37.2151)	−0.0209
Pensions 2003–8				
Absence (average, grades 1 and 4)	−330.0940*** (71.2631)	−0.0103	−396.0887** (169.3640)	−0.0124

Notes: Each panel reports the coefficient associated with the total days of absence (average over grades 1 and 4) in separate regressions where the dependent variable is indicated in the left column. Controls include gender, born out of wedlock, twin birth, born in hospital, mother employed at the time of birth and sets of fixed effects for the year and month of birth, year and month interactions. Specifications also controls for class size, the lowest and highest grade taught to students in the same classroom and length of the school year in weeks, averaged across grades 1 and 4. Models also control for school and teacher fixed effects, and a dummy if the individual changes school or teacher between grades 1 and 4. The OLS models further control for parent's year of birth and family's socio-economic status based on the first-digit HISCO code of the father. Number of observations: more than *Folkskola* 5,976, employment 1960 7,434, employment 1970 5,976, income 1970 5,886, pensions 4,772. Standard errors clustered at the parish level are given in parentheses. Significance: $^{*}p \leq 0.1$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$.

Next, we turn to the effects of absence on labour market outcomes. Impacts on employment are not statistically significant, but the point estimates suggest that impacts on employment may grow slightly with age, from a 0.5% reduction in employment at ages 25–30 to a 1.3% reduction ten years later. As shown in row 3 of Table 3, a one-day increase in the average number of days of absence in elementary school leads to an SEK 80 reduction in earnings in 1970 and an SEK 396 reduction in pensions;[11] both estimates are significant at the 5% level. Relative to the means of these variables, the impacts are fairly similar in magnitude to each other, between 1%–2% reduction in income. To put these magnitudes in perspective, consider that Fredriksson *et al.* (2013) estimated an earnings increase of 1.2% at age 27–42 when class size in Swedish schools is reduced by one student, Chetty *et al.* (2014) estimated an earnings rise of 1.3% at age 28 from a 1-SD improvement in teacher quality in the United States, while Carrell *et al.* (2018) drew on data for a Florida county and showed that exposure to a disruptive peer during elementary school reduces earnings at age 26 by 3%–4%.

As with the short-term outcomes, we explore whether the effects of school absence on long-term outcomes are non-linear, but find no strong evidence that this is the case (Online Appendix

[11] Online Appendix Table D1 shows that pension results are robust when controlling for the receipt of a widow pension.

Figure G1). We also find no strong evidence that the effect of absence depends on whether absence happened in grade 1 or grade 4 (Online Appendix Table G2).

Overall, 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 for both genders. In terms of magnitude, the effects are larger than what is implied by the short-run effects of absence on school performance.[12] Moreover, the effects become slightly larger and more significant when measured further into the life cycle. Taken together, these results are consistent with a dynamic model of skill accumulation, where skill deficits resulting from absences in elementary school translate into lower skill levels progressively accumulating over the life cycle. These results also underline the importance of measuring impacts of absence at various points in the life cycle.

3.3. *Robustness Checks*

The interpretation of the above results as causal effects of absences on long-term outcomes relies on the assumption that school absences are uncorrelated with unobservable determinants of long-term outcomes that vary across siblings. There are several threats to the validity of this assumption. First, long-term outcomes could be affected by child-specific unobserved endowments, which may also affect children's likelihood to miss school. For example, one sibling may be born frailer, have lower grit and work ethics than the other, which would make her more likely to both be absent and have worse adult outcomes than her sibling. Second, long-term outcomes could be affected by idiosyncratic, child-specific shocks that directly affect the incidence of absence and have long-lasting effects on adult outcomes (over and beyond the absence). An obvious example are health shocks during elementary school, which make one sibling (but not the other) particularly sick one year and have long-lasting effects on his/her adult outcomes. Other types of relevant shocks may include changes to the institutional and/or economic environments over time, which affect siblings differently because they are born in different years.

Below we discuss a series of empirical exercises to help us gauge the likely salience of these different threats to the validity of our research design. First, we re-estimate the short-term effect of absence on educational achievement using the sibling FE strategy and compare it to the individual FE estimates of Table 2. As reported in Online Appendix Table G4, the two estimators yield very similar estimates of the impact of absence on achievement in elementary school. This suggests that child-specific unobservable determinants of educational achievement that vary between siblings are unlikely to be correlated with absences. While this does not guarantee that the same is true of unobservable determinants of long-term outcomes, it does reinforce the message that selection into absence based on unobservables is unlikely to be strong.

We then estimate the effects of grade-4 absence on grade-1 performance. If sibling FEs fully account for unobserved child-specific time-invariant factors, future absence should have no power in predicting grade-1 performance. Online Appendix Table G5 shows that, while grade-4 absence and grade-1 performance are negatively correlated in the OLS model, the grade-4 absence coefficient drops and becomes insignificant in the sibling FE model.

[12] Specifically, the (sibling FE) effect of school performance, measured as average grade points in grades 1 and 4, on secondary schooling completion is about 0.07 (Online Appendix Table G3). Multiplying this estimate by the estimated effect of absence on school performance of −0.0045 (Table 2), we would expect an effect of absence on enrolment of −0.0003 in the sibling FE specification. In contrast, our direct sibling FE estimate in Table 3 (−0.005) is considerably larger than the indirect estimate.

The next exercise uses the idea that, if the sibling FE specification does not control for an important unobservable (fixed or time-varying) determinant of long-term outcomes, the sibling FE estimate of grade-4 absence on long-term outcomes would change through the inclusion of grade-1 performance, to the extent that grade-1 performance captures these unobservables (e.g., idiosyncratic work ethic or health shock occurring before grade 1). The results are reported in the second and third columns of Table 4, and show that the estimated effect of grade-4 absence on long-term outcomes remains virtually unchanged through the inclusion of grade-1 performance.

We then conduct two exercises to assess the possibility that individual-specific health shocks during childhood confound our long-term estimates. First, we test whether absence due to sickness affects long-term outcomes *differently* from absence due to other reasons. If health shocks simultaneously determined absence and long-term outcomes, we would likely find that absence due to sickness is more detrimental than absence due to other reasons. We test this hypothesis by re-estimating our model, also controlling for the average number of non-sickness days across grades 1 and 4. We report the coefficients for any absence in the fourth column of Table 4 (see Online Appendix Table G7 for non-sickness absence coefficients and Online Appendix Table F7 for the short-term estimations by reason of absence). Across all long-run outcomes, we cannot reject that the coefficient on the average days of non-sickness absence is equal to the coefficient on the average days of (any) absence, which suggests that the long-term effects of absence in Table 3 are more likely to be driven by the loss of instructional time rather than by a health shock.

Our second test for confounding health shocks looks directly at whether school absences have an effect on long-term health, as measured by mortality. To do so, we re-estimate our sibling FE model, this time using indicators for mortality at three time points (1960, 1970 and 2003) as outcomes. We cannot reject that the coefficient on absence in any of these regressions is zero (see panel B of Table 4). Again, this suggests that the effect of absence we identify through the sibling FE strategy is likely to capture the effect of losing instructional time rather than the effect of a health shock associated with the absence.

Given that the labour market was gender segmented, we also test whether sibling sex composition matters. Column (5) of Table 4 reports estimates for the sample of families with same-sex siblings, which are very similar to our main results and suggest that low female labour force participation is unlikely to bias the analysis using the entire sample. The last two columns of Table 4 show that the baseline results are also robust to controlling for siblings' differential exposure to the instructional time school reform and to economic environments at birth.

3.4. *Bounds*

The above robustness checks all suggest that the identifying assumption underlying the sibling FE strategy is unlikely to be violated. However, given that it does remain an untestable assumption, we conclude our analysis by bounding our estimates, using the approach of Oster (2019), which is also implemented in Liu *et al.* (2021). As we describe in Online Appendix H, this approach exploits observables in order to gauge how unobservables may affect the estimates, and is therefore only helpful to the extent that the selection on observables is informative of the selection on unobservables one might suspect. As our empirical strategy controls for a breadth of factors that are likely to affect absence and long-term outcomes, the bounding approach seems useful here.

Column (1) ('short regression') of Table 5 reports the coefficients associated with absence in a regression without any controls. Column (2) ('intermediate regression') replicates the sibling FE

Table 4. *Robustness of the Effect of Absence on Long-Term Outcomes.*

Reported coefficient on:	(1) Baseline results (Table 3)	(2) (*i*) Grade-4 effect with and without controlling for grade-1 performance	(3)	(4) (*ii*) Differentiating for reason of absence	(5) (*iii*) Same-sex siblings	(6) (*iv*) Contr. for compulsory schooling	(7) (*v*) Contr. for local conditions at birth
	Avg., grades 1 and 4	Grade-4 absence	Grade-4 absence	Avg., grades1 and 4	Avg., grades 1 and 4	Avg., grades 1 and 4	Avg., grades 1 and 4
Panel A: estimates of the effect of absence in different sibling fixed effect specifications							
More than *Folkskola* (1 = yes)	−0.0005	−0.0010	−0.0011	−0.0006	−0.0011	−0.0005	−0.0004
	(0.0009)	(0.0008)	(0.0008)	(0.0007)	(0.0016)	(0.0009)	(0.0012)
Employment 1960 (1 = yes)	−0.0007	0.0016	0.0016	−0.0005	−0.0023	−0.0007	−0.0016
	(0.0013)	(0.0011)	(0.0011)	(0.0010)	(0.0019)	(0.0013)	(0.0014)
Employment 1970 (1 = yes)	−0.0020	−0.0016	−0.0016	−0.0015	−0.0048	−0.0019	−0.0039***
	(0.0014)	(0.0010)	(0.0010)	(0.0012)	(0.0030)	(0.0014)	(0.0013)
Labour market income 1970	−80.7**	−61.1	−62.1	−50.0	−117.4	−81.1**	−85.9*
	(37.2)	(45.4)	(45.2)	(32.7)	(85.14)	(37.3)	(45.9)
Pensions 2003–8	−396.1**	−284.9**	−248.2**	−319.1***	−689.2*	−395.8**	−572.1***
	(196.4)	(119.4)	(105.0)	(121.5)	(405.4)	(169.2)	(199.8)
Panel B: additional results—effects of absence (average, grades 1 and 4) on mortality							
Passed away before 1960 (1 = yes)	0.00001						
	(0.00030)						
Passed away before 1970 (1 = yes)	0.00036						
	(0.00048)						
Passed away before 2003 (1 = yes)	0.00056						
	(0.00078)						

Notes: Each panel reports the coefficient associated with the total days of absence (average over grades 1 and 4) in separate regressions where the dependent variable is indicated in the left column. Controls as reported in Table 3. Panel A shows estimates in different sibling fixed effect specifications: (*i*) grade-4 effect with and without controlling for grade-1 performance (see Online Appendix Table G6 for grade-1 performance coefficients), (*ii*) differentiating for different reasons of absence (see Online Appendix Table G7 for non-sickness absence coefficients), (*iii*) estimates using same-sex siblings, (*iv*) controlling for school reforms and (*v*) controlling for local conditions at birth (poverty rate, taxable income per capita, economic development (the year-to-year change in the taxable income per capita), all assessed at the time of birth). Panel B shows estimates on absence on mortality by 1960, 1970 and 2003. Standard errors clustered at the parish level are given in parentheses. Significance: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table 5. *Oster Bounds for the Long-Term Effects of Absence in School.*

	(1) Coefficient of absence	(2)	(3) Selection bias	(4)
Dependent variable	Short regression	Intermediate regression	Same direction	Opposite direction
More than Folkskola *(1 = yes)*				
Absence (average, grades 1 and 4)	−0.0009* (0.0005) [0.00]	−0.0005 (0.0009) [0.41]	−0.73850	−0.00014
Employment 1960 (1 = yes)				
Absence (average, grades 1 and 4)	−0.0013 (0.0008) [0.00]	−0.0007 (0.0013) [0.60]	0.00306	−0.00177
Employment 1970 (1 = yes)				
Absence (average, grades 1 and 4)	−0.0008 (0.0008) [0.00]	−0.0020 (0.0014) [0.50]	−0.04959	−0.00155
Labour market income 1970				
Absence (average, grades 1 and 4)	−65.4768*** (22.3882) [0.00]	−80.7042** (37.2151) [0.64]	−107.88081	−73.15848
Pensions 2003–8				
Absence (average, grades 1 and 4)	−384.5674*** (81.8270) [0.00]	−396.0887** (169.3640) [0.59]	−90.55889	−423.49521

Notes: Column (1) gives the coefficient associated with absence in the short regression where the outcome variable (stated on the left) is regressed on absence and an intercept (without any controls), denoted by $\dot{\beta}$ in Online Appendix H. The intermediate regression in column (2) reiterates the sibling FE specification estimates of the coefficient associated with absence (reported in Table 3 and corresponding to $\tilde{\beta}$ in the model). Columns (3) and (4) report the bounds (β^*) when selection on unobservables is of the same magnitude as selection on observables and goes in the same direction ($\delta = 1$) and the opposite direction ($\delta = -1$), respectively. The R^2 of the short and intermediate models are reported in brackets in columns (1) and (2), respectively. To calculate the bounds, we use the Stata ado-file `psacalc` provided online by Emily Oster. All errors are our own responsibility. Number of observations as in Table 3. Standard errors clustered at the parish level are given in parentheses.

estimates from Table 3. Column (3) (column (4)) reports the bound under the assumption that the unobservables move the point estimate in the same (opposite) direction and by the same magnitude as the introduction of the observables does. Despite the R^2 of the intermediate regression being higher than 0.5 for all outcomes, there is remarkably little change between columns (1) and (2). This modest influence of selection on observables is reflected in the estimated bounds: all bounds are negative, with the exception of the bound for the effect of absence on 1960 employment in column (3). Thus, even under the extreme scenario that our sibling FE model fails to account for as much selection on unobservables as it does account for selection on observables, we cannot reject there would be a negative effect of absence on long-term outcomes.[13]

4. Conclusion

School absences are 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

[13] Bounds for short-term effects in Online Appendix Table H1 underline this conclusion.

providing causal evidence of the impact of student absence on academic performance focus on the United States. The 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 and negative impact on student performance: increasing total absence in one grade by ten days leads to a reduction in grade point average of 4.4% of an SD, an effect of comparable magnitude to that found in the United States. For men, this immediate impact on school performance spills over into 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 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. Together, the short- and long-term effects of absence suggest that a key mechanism underlying these results is the effect of instructional time losses on early levels of skills, which accumulate over the life cycle and eventually create non-negligible income penalties.

Our research starts filling an important gap in the evidence base on the long-term impacts of school absence and thereby informs policy discussions about high rates of absences around the world. 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 COVID-19 pandemic. But, as school absenteeism becomes 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.

Institute for Fiscal Studies, London, UK & IZA—Institute of Labor Economics, Bonn, Germany
Heinrich Heine University Düsseldorf, Germany & IZA—Institute of Labor Economics, Bonn, Germany
University of Duisburg-Essen, Germany & IZA—Institute of Labor Economics, Bonn, Germany
Lund University, Sweden & Research Institute of Industrial Economics (IFN), Stockholm, Sweden

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Agüero, J. and Beleche, T. (2013). 'Test-Mex: Estimating the effects of school year length on student performance in Mexico', *Journal of Development Economics*, vol. 103(C), pp. 353–61.

Altonji, J.G., Kahn, L.B. and Speer, J.D. (2016). 'Cashier or consultant? Entry labor market conditions, field of study, and career success', *Journal of Labor Economics*, vol. 34(S1), pp. S361–401.

Aucejo, E.M. and Romano, T.F. (2016). 'Assessing the effect of school days and absences on test score performance', *Economics of Education Review*, vol. 55, pp. 70–87.

Bergh, A. (2014). *Sweden and the Revival of the Capitalist Welfare State*, Cheltenham, UK: Edward Elgar Publishing.

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*, vol. 15(5), pp. 1101–57.

Bhuller, M., Mogstad, M. and Salvanes, K.G. (2017). 'Life-cycle earnings, education premiums, and internal rates of return', *Journal of Labor Economics*, vol. 35(4), pp. 993–1030.

Bond, T.N. and Lang, K. (2013). 'The evolution of the black-white test score gap in grades K-3: The fragility of results', *The Review of Economics and Statistics*, vol. 95(5), pp. 1468–79.

Carlsson, M., Dahl, G., Öckert, B. and Rooth, D.O. (2015). 'The effect of schooling on cognitive skills', *The Review of Economics and Statistics*, vol. 97(3), pp. 533–47.

Carrell, S.E., Hoekstra, M. and Kuka, E. (2018). 'The long-run effects of disruptive peers', *American Economic Review*, vol. 108(11), pp. 3377–415.

Chetty, R., Friedman, J.N. and Rockoff, J.E. (2014). 'Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood', *American Economic Review*, vol. 104(9), pp. 2633–79.

Erixon, L. (2008). 'The Swedish third way: An assessment of the performance and validity of the Rehn–Meidner model', *Cambridge Journal of Economics*, vol. 32(3), pp. 367–93.

Federation of Swedish Genealogical Societies. (2014). 'Swedish death index 1901–2013', Data files, Farsta, Sweden: Federation of Swedish Genealogical Societies.

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*, vol. 18(6), pp. 2776–823.

Fitzpatrick, M., Grissmer, D. and Hastedt, S. (2011). 'What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment', *Economics of Education Review*, vol. 30(2), pp. 269–79.

Fredriksson, P., Öckert, B. and Oosterbeek, H. (2013). 'Long-term effects of class size', *The Quarterly Journal of Economics*, vol. 128(1), pp. 249–85.

Galama, T.J., Lleras-Muney, A. and van Kippersluis, H. (2018). 'The effect of education on health and mortality: A review of experimental and quasi-experimental evidence', *Oxford Research Encyclopedia of Economics and Finance*, Oxford University Press, doi: 10.1093/acrefore/9780190625979.013.7.

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

Genda, Y., Kondo, A. and Ohta, S. (2010). 'Long-term effects of a recession at labor market entry in Japan and the United States', *Journal of Human Resources*, vol. 45(1), pp. 157–96.

Goodman, J. (2014). 'Flaking out: Student absences and snow days as disruptions of instructional time', Working Paper 20221, National Bureau of Economic Research.

Hansen, B. (2011). 'School year length and student performance: Quasi-experimental evidence', Unpublished manuscript, University of Oregon.

Leuven, E., Lindahl, M., Oosterbeek, H. and Webbink, D. (2010). 'Expanding schooling opportunities for 4-year-olds', *Economics of Education Review*, vol. 29(3), pp. 319–28.

Liu, J., Lee, M. and Gershenson, S. (2021). 'The short- and long-run impacts of secondary school absences', *Journal of Public Economics*, vol. 199, 104441.

Marcotte, D. and Hansen, B. (2010). 'Time for school?', *Education Next*, vol. 10(1), pp. 53–9.

Marcotte, D. and Hemelt, S. (2008). 'Unscheduled school closings and student performance', *Education Finance and Policy*, vol. 3(3), pp. 316–38.

Oreopoulos, P., Von Wachter, T. and Heisz, A. (2012). 'The short- and long-term career effects of graduating in a recession', *American Economic Journal: Applied Economics*, vol. 4(1), pp. 1–29.

Oster, E. (2019). 'Unobservable selection and coefficient stability: Theory and evidence', *Journal of Business & Economic Statistics*, vol. 37(2), pp. 187–204.

Paulsson, E. (1946). *Om folkskoleväsendets tillstånd och utveckling i Sverige under 1920-och 1930-talen (till omkring år 1938)*, Jönköping, Sweden: 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', Economic Journal, vol. 117(523), pp. 1216–42.

SCB. (1962). 'Population and housing census 1960', Data files, Stockholm, Sweden: National Bureau of Statistics (SCB).

SCB. (1972). 'Population and housing census 1970', Data files, Stockholm, Sweden: National Bureau of Statistics (SCB).

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*, vol. 22(1), pp. 133–48.

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