

A researcher's guide to the Swedish compulsory school reform*

Helena Holmlund

IFAU – Institute for the Evaluation of Labour Market and Education Policy
UCLS
Ifo Institute

Abstract

This paper demonstrates how a natural experiment in education can be used to estimate causal effects. The Swedish compulsory school reform extended basic education gradually across cohorts and municipalities, allowing for a difference-in-differences analysis. The paper summarizes the literature using this reform and shows that it provided individuals from low socio-economic backgrounds with better opportunities in life. Not only did they attain higher levels of education – they also earned higher earnings, were less likely to participate in crime, and more likely to run for office.

JEL-codes: I24; I26; I28

**I wish to thank two anonymous referees, Anders Björklund, Mikael Lindahl, Olmo Silva and participants at the CEE group meeting at LSE for helpful comments. Thanks also to Anders Björklund, Valter Hultén and Mikael Lindahl for help with the coding of the reform. Parts of the research has been financed by the Swedish Council for Working Life and Social Research, and their support is gratefully acknowledged. IFAU, Kyrkogårdsgatan 6, 751 20 Uppsala, Sweden. Email: helena.holmlund@ifau.uu.se.*

1. Introduction

In the 25 years following the Second World War, many western European countries undertook major educational reforms with the main purpose of extending compulsory education. The Nordic countries, continental Europe and the United Kingdom, with different traditions of education policy, were all part of this widespread expansion (Viarengo 2007). The strong economic growth in the post war era created a demand for a higher skilled workforce, and in some countries, for example Sweden, there was a strong push for reforming the education system in order to increase equality of opportunity. The European experience was also a reflection of an earlier development in the United States, where compulsory school attendance and child labour laws were enacted throughout the states in the early decades of the 20th century. These compulsory schooling reforms have spurred an enormous interest in applied economics; reforms that because of their design offer the promise of estimating causal effects of extending basic education on a range of outcomes such as earnings, health, crime, and intergenerational effects on the education of offspring in the next generation.

Sweden extended compulsory education gradually across the country, starting in the late 1940s. The reform was implemented in different municipalities at different points in time, meaning that for a given birth cohort some individuals went through the old two-tier selective system where basic education ended after 7 or 8 years, and others went through the new school, comprising of one or two more years in a comprehensive system. Similar gradual expansion paths were also adopted in the other Nordic countries, and the design of these reforms has resulted in a large number of studies that exploit variation across regions and over time as a source of quasi-random variation in both length of compulsory schooling and/or educational tracking (seminal papers are Meghir and Palme 2005 for Sweden, Black et al. 2005 for Norway, and Pekkarinen et al. 2006 for Finland).¹ Closely related to these papers is the U.S. literature, which with a similar methodological approach studies compulsory school leaving ages across U.S. states (see for example Lochner and Moretti 2004 and Lleras-Muney 2005).

Today, 15 years after Meghir and Palme's paper was first published, there is a large body of research based on the Swedish compulsory school reform; the combined use of Swedish register data and difference-in-differences analysis exploiting the gradual nature of reform implementation has given rise to many good publications. This paper summarizes the existing papers up to date and discusses the insights from the literature in light of the political background of the reform. In addition, the paper offers a documentation of the reform data collection previously presented in Holmlund (2007), which has been to the benefit of many of the papers cited here. More specifically, I discuss how information on reform implementation can be linked to different sources of data, I present balancing tests to examine if reform exposure is conditionally correlated with individuals' observed characteristics and report estimates of the effect of the reform on educational outcomes. Using two independent measures of reform assignment, I also run IV regressions that take into account measurement error bias and bound the reform "first stage" estimates.

So, what have we learned? First, as I show in the empirical part of this paper, the difference-in-difference approach is successful since both balancing tests and tests for pre-reform parallel trends suggest that the underlying assumption of parallel trends is satisfied. Second, the reform had larger effects on educational attainment among individuals from lower SES backgrounds, thus contributing to intergenerational mobility. Finally, summarizing the literature today, I conclude that an impressive number of good publications have emerged from combining register data and the reform. We have learned that the extension of basic education has had positive effects on cognitive skills – but that non-cognitive characteristics among children from high SES backgrounds might have been negatively affected. We can also conclude that the reform reduced criminal involvement, and that individuals from working class backgrounds became more likely to engage politically by running for office. Studies on mortality and health show diverging results, but the overall impression is that there are no effects on health outcomes. However, the reform seems to have affected financial decision-making. Lastly, the reform has proven to affect outcomes in the next generation: there are spillover effects to the skills of targeted individuals' children.

¹ The Danish school reform was not subject to gradual implementation and as such has not been used as widely to estimate causal effects. See Arendt (2008) for a description of Danish reforms and an application.

The remainder of the paper unfolds as follows: section 2 presents a general discussion on estimating returns to education, section 3 covers institutional details regarding the Swedish compulsory school reform, and sections 4 and 5 detail reform assignment and data sources. Section 6 presents balancing tests and estimates of reform effects on educational outcomes. Section 7 summarizes the literature using the Swedish reform up to date, and finally, section 8 offers conclusions. For more detail, Appendix A contains a documentation of sources used to determine reform status.

2. Returns to education

There is a long tradition in empirical labour economics of studying returns to education. The early literature focused on the pecuniary returns – the percentage wage gain from one more year of schooling. Later, this literature was extended to focus on non-pecuniary returns to education, such as effects of education on health, crime and fertility. These studies all face the same challenge, that is, how to control for unobserved ability or other unobserved factors that are correlated with education and also with the outcome of interest. For example, is the wage premium from schooling truly an effect of education in itself, or does it purely reflect the fact that more able or more motivated workers, earning higher wages, also choose a higher level of education? Moreover, estimating the effect of education on health, how do we take into account that individuals in poor health might not find it worthwhile investing in education because their health condition implies lower returns to their investment (an example of reverse causality)? Health and education can also be correlated because of discount rate bias: individuals with a high discount rate might invest neither in human capital nor health. And similarly, with fertility as with many other potential outcomes, how can we as researchers control for unobserved preferences that jointly determine both education and fertility outcomes?

The remedy in many studies of these issues has been either to control for ability by using samples of identical twins (Ashenfelter and Krueger 1994, Behrman and Rosenzweig 1999), or to make use of some exogenous source of variation in education, typically in the form of a natural experiment. Natural experiments offer variation in some treatment, in this case compulsory education legislation, for individuals that otherwise can be assumed to be identical; natural experiments allow us to come around the problem of education being correlated with individual characteristics such as ability or motivation. There are a number of well-known examples, apart from the Nordic papers cited above, and the Swedish papers cited in section 7. Angrist and Krueger (1991) use the fact that a fixed school-leaving age in the US allows students to drop out from school earlier if they are born early in the year, that is, the length of compulsory education varies with month of birth.² Acemoglu and Angrist (2000), Lochner and Moretti (2004), Lleras-Muney (2005), Oreopoulos et al. (2006) all use variation across US states in compulsory schooling laws: variation across states is introduced by both compulsory attendance and child labour laws. Currie and Moretti (2003) account for endogeneity of schooling by using variation induced by college openings. Chevalier (2004) studies a reform in the UK and Maurin and McNally (2005) use variation in schooling introduced by the 1968 revolts in Paris. The outcomes in the studies mentioned above all range from the pecuniary return to education, to effects on mortality, crime, birth outcomes and the education of the children (the intergenerational effect of education).

3. The Swedish reform

The Swedish educational reform is carefully described in the work by Marklund (1980, 1981). Detailed information can also be found in a report by the National Board of Education (1960). The following brief description builds on these sources, which are recommended for further details on the topic.

Prior to the school reform, pupils in Sweden went through grades 1 to 4 or 1 to 6 in a common school (*folkskolan*). In either fourth or sixth grade, more able students were selected (based on past performance) for the five or three/four-year

² Later research has invalidated this approach by showing that month of birth in itself is directly related to educational outcomes, through e.g. maturity at school starting-age, or through the benefits of relative age in the classroom (see e.g. Bound and Jaeger 1996 for a critical discussion).

long junior-secondary school (*realskolan*). Remaining students stayed in the common school until compulsory education was completed. In most cases, compulsory education comprised seven years, but in some municipalities, mainly the big cities, the minimum was eight years. The system resembled the traditional European model with early selection, parallel school forms and a small tertiary sector (Erikson and Jonsson 1996).

In 1946, the social-democratic government appointed a parliamentary committee (*1946 års skolkommission*) which was given the task to analyze the Swedish school system and to develop proposals and guiding principles for a non-selective compulsory school. The main purpose with such a change was to postpone the tracking decision to higher grades, in an effort to increase equality of opportunity.³ Two years after the appointment of the committee, in 1948, the committee released its proposals. The main suggestion was to introduce a nine-year compulsory school, where pupils were kept together in common classes longer than in the earlier school system. As a compromise between the opponents of early tracking and its advocates, the committee proposed tracking in 9th grade; pupils would follow either a vocational track, a general track, or a theoretical track preparing for upper-secondary school. The 9th grade streaming was later abandoned in favor of a completely comprehensive system.

Erikson and Jonsson (1996) argue that more than in most other Western countries, school reforms in Sweden have been characterized by the specific aim of reducing social and educational inequalities. Early selection was considered a hurdle for children from low socioeconomic backgrounds to access secondary education, and with a comprehensive system the idea was to provide equal opportunities for all children, regardless of family background. Naturally, the reform also had its critics and the question of tracking became the key controversy around the reform, with the right-wing party in opposition of late selection.

To evaluate the appropriateness and whether the proposed nine-year comprehensive school would serve its purpose, in 1949 the committee suggested that an “experiment” would take place, where during an assessment period some municipalities and schools would implement the new school system such that the results could be scrutinized before further decisions were made.

The assessment programme came to start in 1949/1950, this year under the surveillance of the parliamentary committee. In 1950, the Swedish parliament committed to the introduction of a nine-year comprehensive school and approved of the idea of a trial period at the outset of the reform. When the formal decision was made in 1950, the National Board of Education (*Skolöverstyrelsen*) took over the administration of the reform.

The new comprehensive school was to be introduced throughout a whole municipality, or in certain schools within a municipality. Following the 1948 proposal of the parliamentary committee, a number of municipalities had declared interest in reforming their comprehensive schools. For this reason, 264 municipalities (out of around 1000) were asked if they were willing to introduce the nine-year school immediately or within a few years. The municipalities that were approached had either shown interest in the reform or expanded their junior secondary school to four years. 144 municipalities showed interest in the reform. 14 municipalities were selected for the first year of the assessment (1949/50), all of those were required to have an eight-year comprehensive school already.

The following years, the National Board of Education continued with the implementation of the reform. Year by year, more municipalities joined the reform assessment programme. Municipalities that wanted to take part in the reform were asked to report on their population growth, on the local demand for education, tax revenues and local school situation. For example, the availability of teachers, the number of required teachers for the nine-year comprehensive school, and the available school premises were explored. The National Board of Education took these municipality characteristics into account when deciding on their participation. In general, implementation of the reform started in grades 1 and 5, the following year covering grades 1, 2, 5 and 6 and so on. From 1958 the reform was introduced in grades 1–5 already from the starting year.

³ *The large baby boom cohorts that passed through the education system during this period, and higher demand for junior secondary schooling overall, are also likely drivers behind reforming the education system.*

Apart from extending compulsory education from seven (or in some cases eight) years to nine years, and to postpone tracking, the educational reform was also pedagogical and affected the curriculum somewhat. The main change of the curriculum was that English was introduced in 5th grade in the new comprehensive school, while this was not necessarily a compulsory subject in the old school system. The school starting age was set at the year the child turned seven in both the old system and the new comprehensive school.

The assessment period was also accompanied by financial support to families and to municipalities that implemented the reform. A universal child allowance was introduced in 1948 and implied support for children until the age of 16. In reform municipalities, a means-tested scholarship compensated families for foregone earnings from keeping their children longer in school. Municipalities were compensated for the increased costs following the expansion of education. The state provided funds targeted at the new comprehensive school, one example is complete funding of teacher salaries for grades 7–9, in the years 1952–1955.

In 1962, the parliament came to a final decision to permanently introduce the nine-year school throughout the country. At this point, the implementation came to be a matter for each municipality; by 1969 they were obliged to have the new comprehensive school running. Since the timing was much in the hands of each municipality, the implementation was far from a randomized experiment, but nevertheless provides a source of variation in schooling laws that may be fruitfully explored by the empirical researcher.

4. Linking individuals to treatment

Since the educational reform provides a potential source of quasi-random variation in education, I take a closer look at the available data. For the quantitative researcher, knowledge about which municipalities implemented the reform, and which birth cohorts were affected, is of particular importance. Below, I list three different data sources available to study reform effects on individual outcomes.

1. The IS data

It is possible to use the IS (individual statistics) data, from the Institute of Education at Gothenburg University (Härnqvist 2000). The data stem from surveys, conducted in 6th grade, of around 10 percent of the cohorts born in 1948 and 1953. When these data were collected, information on type of school (the old *folkskola* or the new nine-year comprehensive school) that each individual attended was recorded, based on information provided by the local school. Register information on adult earnings and other register-based information can be matched to individuals in the data. This is the data set explored in Meghir and Palme's (2005) work on the Swedish compulsory school reform.

2. The Swedish Level of Living Survey

The Swedish Level of Living Surveys, based on random samples of the Swedish adult population, have been conducted in 1968, 1974, 1981, 1991, 2000 and 2010 (Erikson and Åberg 1984). The surveys ask specifically whether an individual went through the old system or the new nine-year school. These data have been used by Jasmina Spasojevic (2010) in her work on the effects of education on health.

3. Register data from Statistics Sweden

The Swedish administrative registers do not contain information on whether individuals in the affected cohorts went through the old or the new school system. With help from other sources (described in Appendix A) it is however possible to deduct when and for which grades each municipality introduced the new comprehensive school and based on this information one can assign reform status to the individuals in a data set extracted from registers. Censuses can be used to track in which municipality an individual lived at the time of compulsory education. With

this information it is possible to attach a reform indicator to each individual based on year of birth and municipality of residence, maintaining the assumption that individuals are in the right grade according to their age. In some cases, it is also necessary with more detailed information on in which parish or school district the individual went to school, since the reform was sometimes introduced in parts of a municipality in different years. Any given dataset with information on birth year and municipality/parish of residence can assign reform participation to the cohorts that were subject to the education reform.

There are two possible ways to construct a reform coding that can be matched to individual-level register data. In the remainder of the paper I will label them coding 1 (based on documentation) and coding 2 (deduced from register data).

The sources of information necessary to construct coding 1 of the reform implementation are the following:

- Marklund (1981) and the National Board of Education (1953–1962). These sources document the assessment programme when the reform was gradually introduced across the country, and they include lists of which municipalities implemented the reform each year. In the latter publication it is also possible to see which grades that were affected in a particular municipality. These sources cover the assessment period and only allow coding of the cohorts born 1938–1949.
- The Educational Bureau (Undervisningsbyrån) (1960–1964) and Statistics Sweden (1968–1969). From municipality-level tables of the number of pupils in each grade in the old and new school system, it is possible to deduct when the reform was implemented, and the remaining cohorts can be coded.

Register data sets with large sample sizes allow for an alternative procedure to assign reform status. Coding 2 is simply obtained by splitting the sample, using the first part of the sample to identify treatment status by empirically observing when the minimum level of education (by municipality/birth year) jumps up from *folkskola* (the old compulsory minimum) to *grundskola* (the new minimum), and using the second part of the sample to estimate effects of treatment on outcomes.

Appendix A explains in detail how the different sources have been used to create coding 1 and coding 2, and also highlights some of the difficulties relating to the coding of some municipalities, where the reform was not implemented universally at one point in time. In the remainder of this paper I discuss and demonstrate the reform, and explore several important data issues, using a data set compiled from Swedish registers. Coding 1 is available from the author for researchers who wish to use it.

5. Data

The empirical analyses are based on data from Swedish administrative registers, available for researchers through Statistics Sweden. The population of interest is defined as cohorts born in Sweden between 1945 and 1955.⁴ By linking several registers and censuses through personal identifiers, individuals can be linked to their parents, and to information on both family background and long-term outcomes.

Family background. Family background is characterized by father's education level and father's earnings. Father's education level is derived from the 1970 census and is defined as a dummy for high education, which represents any education above compulsory level. Father's earnings is a measure of average earnings over the years 1968, 1971 and 1973, percentile ranked within father's own birth cohort. These variables are used in a balancing test to investigate whether treatment is correlated with pre-determined characteristics.

⁴ This population is chosen because the birth years span the vast part of the new school expansion, and because it is possible to assign reform status to these cohorts.

Educational outcomes. I study educational outcomes through information on highest completed level of education in the education register from 1995. Highest level of education has been translated into years of education by assigning the expected number of years to each level.⁵ I also study the probability of attaining any post-compulsory education beyond the new compulsory minimum. To arrive at this measure, I use observed levels in the education registers, which implies education at secondary or tertiary level.

Assigning reform status. Reform status is assigned to individuals based on birth year and municipality of residence. The cleanest way to assign treatment is to use information on pre-treatment location, which in my data is available as municipality of birth, derived from the parish where the child was registered at birth. However, for cohorts born until 1946, the parish of birth that was reported refers to the location of the hospital in which they were born (Skatteverket 2007). At this time, most births did take place out of the home, and a majority of all municipalities did not have their own maternity ward. Thus, the information on parish of birth cannot be used to identify treatment status for early cohorts in the sample.

The alternative (used in this paper) is to extract census information on home municipality in 1960 and 1965, approximately at age 10–15. This arguably assigns reform status to individuals based on where they lived while going to school but could be considered as endogenous to reform exposure if families move in response to the reform.

After excluding individuals with missing observations on either education or municipality of residence in 1960/1965, I arrive at 1,182,063 observations. Treatment is identified for 91 percent (using coding 1 based on documentation). Descriptive statistics of the sample are presented in Figure 1.

Table 1. Summary statistics

Variable	Nr of obs	Mean	Std. Dev.
Years of schooling	1,021,996	11.46	2.52
Any post-compulsory schooling	1,021,996	0.75	0.43
Reform exposure	1,021,996	0.48	0.50
Woman	1,021,996	0.49	0.50
High educated father	767,567	0.31	0.46
Income rank father	966,031	52.31	27.86
Birth year	1,021,996	1949.89	3.18

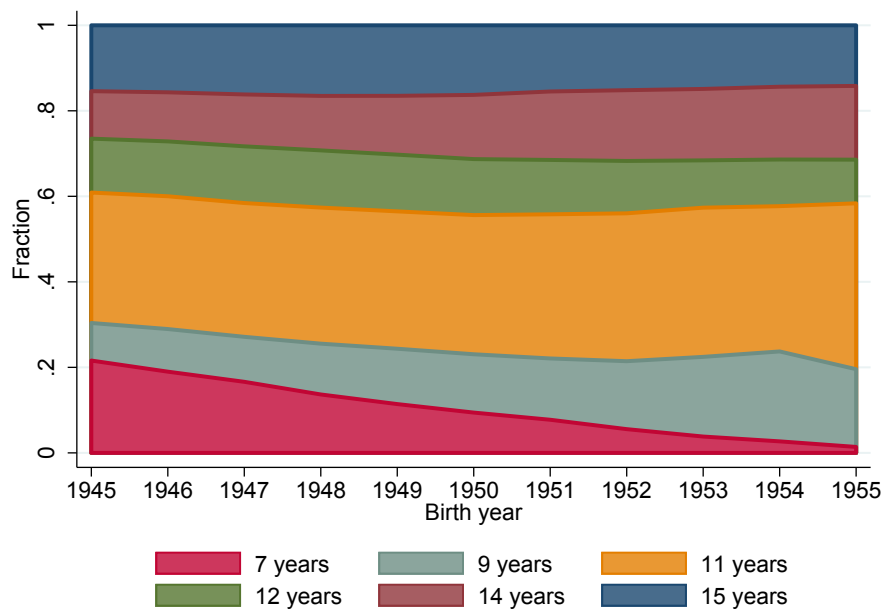
⁵ Fisher et al. (2018) show that deriving years of schooling from education levels can understate reform effects, since some compliers to the reform will be coded with “too much” education using this method. Using alternative data sources, they find that the reform effect on years of schooling is 76 percent larger when taking into account compliers that are not observed when using education levels to identify years of schooling.

6. An analysis of the Swedish compulsory school reform

The compulsory schooling reform affected cohorts whose education levels were on the rise across the board – both at the low and the high end of the distribution. Besides the focus on raising the minimum level, the study grant system for higher education was reformed in 1965, and several new tertiary education institutions were opened in the 1960s and 70s. Accordingly, Figure 1 shows that the number of years of schooling were increasing, both as a consequence of a higher share of individuals attaining 9 years, and higher shares both at upper-secondary (11–12 years) and post-secondary (14 + years) levels.

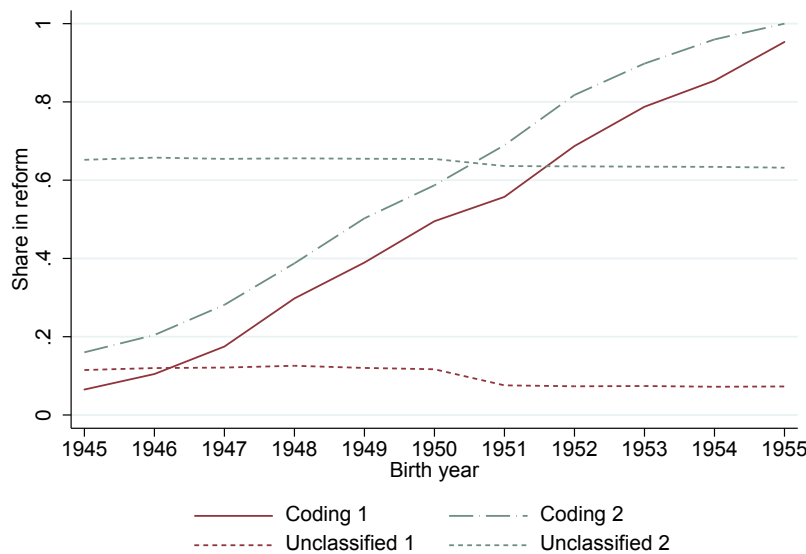
Figure 2 shows the share of individuals treated by the reform, for the two different sources of reform assignment described in section 4. The different coding schemes follow each other closely in terms of the share of individuals in each birth cohort that is affected.⁶ It also shows that for the cohorts depicted in the figure, the increase over time is fairly linear. When reaching the 1955 cohort, almost 100 percent of individuals have been assigned treatment according to the two different coding schemes.

Figure 1. Share of population at different levels of schooling



Note: Years of schooling assigned according to highest level in the 1995 education register. 15 years refers to 15 years or more.

⁶ The large share with unclassified treatment status for coding 2 is explained by the fact that 60 percent of the sample has been used to identify treatment status by municipality/birth year, and 40 percent of the sample is used in estimation.

Figure 2. Share of population treated by the reform

Note: Coding 1 refers to assignment based on documentation, coding 2 refers to reform assignment based on observed minimum levels by cohort and municipality (using 60 percent of the sample to identify treatment status, 40 percent for descriptives and estimation). Reform status has been assigned to individuals based on birth cohort and municipality of residence observed in the censuses 1960 (cohorts 1945–1950) and 1965 (cohorts 1951–1955).

6.1 Reform effects on schooling outcomes

The gradual implementation implies that a staggered difference-in-differences design is a natural starting point for estimating effects of reform exposure on outcomes. Consider the following baseline specification:

$$(1) \quad y_{icm} = \alpha + \beta T_{cm} + \gamma_c + \delta_m + \varepsilon_{icm}$$

where y_{icm} is the outcome of individual i , belonging to cohort c and municipality m . T_{cm} is a treatment indicator that takes the value of 1 if a cohort and municipality is treated, and 0 otherwise. The equation also includes cohort- and municipality-specific fixed effects (γ_c and δ_m). The parameter of interest is β , which gives us the ITT (intention-to-treat) parameter of reform exposure on the outcome of interest. In essence, the method compares the difference in outcomes over time – before and after treatment in treated regions – to the same time difference in untreated regions or in regions treated at a different point in time. The staggered differences-in-differences specification has been common in many empirical applications and reform evaluations but has recently received attention from a methodological point of view: new econometrics papers demonstrate that the method can lead to biased estimates if heterogeneous treatment effects are present. I return to this topic in section 6.6 where I briefly discuss the methodological problems in more detail and perform a sensitivity tests suggested in the recent literature.

The difference-in-differences design relies on the identifying assumption of parallel trends between treated and control units in the absence of any intervention. This assumption cannot be tested, but there are two standard ways to assess the credibility of the assumption. First, a balancing test regressing pre-determined characteristics on treatment and region- and cohort controls sheds light on whether there are differential compositional trends in terms of observed characteristics in treated and control regions. Second, a specification estimating pre-reform effects illustrates whether the assumption holds in the pre-reform period. Before turning to the main effects of the reform on educational outcomes, I present these two analyses in turn.

6.2 Balancing tests

Table 2 begins by presenting balancing tests relating reform exposure to two pre-determined characteristics capturing the socioeconomic background of individuals in the population. The first panel presents results using coding 1, while the second panel corresponds to coding 2. First, columns 1 and 2 present the relationship between treatment and father's education (a dummy for any post-compulsory education). Column 1 presents the baseline specification, while column 2 presents a specification including linear time trends interacted with indicators for implementation year (a simplification of a specification including municipality-specific linear trends). Columns 3 and 4 present the corresponding estimates for father's income percentile. The overall impression is that the difference-in-differences specification is successful in handling any form of sorting into treatment based on observable characteristics; the coefficients are precisely estimated and close to zero.

Table 2. Balancing tests

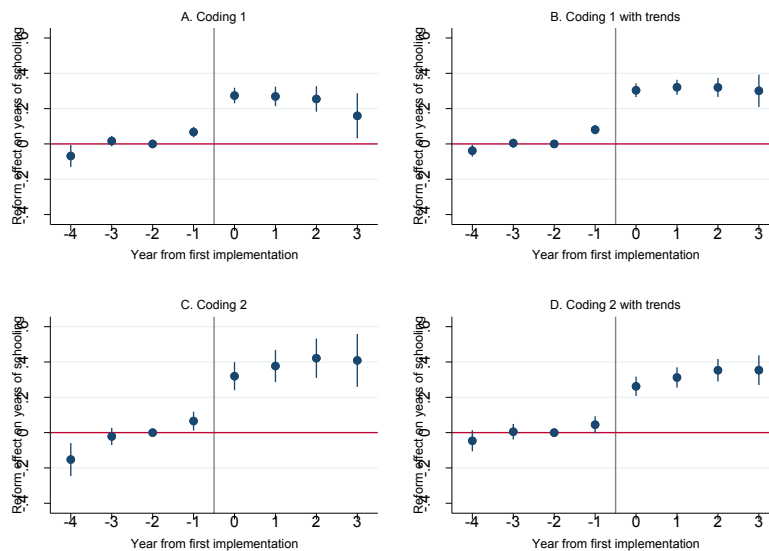
	(1)	(2)	(3)	(4)
	<i>Father high education</i>		<i>Father's income percentile</i>	
A. Balancing test reform coding 1				
Reform exposure coding 1	-0.001 (0.002)	0.002 (0.002)	0.022 (0.113)	0.085 (0.116)
Observations	767,567	767,567	966,031	966,031
R-squared	0.002	0.002	0.000	0.000
Number of municipalities	1,020	1,020	1,020	1,020
Outcome mean	0.310	0.310	52.32	52.32
B. Balancing test reform coding 2				
Reform exposure coding 2	0.005 (0.007)	-0.004 (0.003)	0.587 (0.432)	-0.095 (0.163)
Observations	302,131	302,131	380,059	380,059
R-squared	0.002	0.002	0.000	0.000
Number of municipalities	986	986	986	986
Outcome mean	0.312	0.312	52.38	52.38
Muni f.e.	Yes	Yes	Yes	Yes
Cohort f.e.	Yes	Yes	Yes	Yes
Implement-spec trends	No	Yes	No	Yes

Note: Robust standard errors in parentheses clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The sample size is reduced in panel B since coding 2 is defined for a subset of the population (here 40 percent).

6.3 Event-study analysis

The event-study design expands the difference-in-differences specification by introducing treatment dummies interacted with time since first treatment year. By doing so, we are able to both address the pre-reform parallel trends assumption and depict the time pattern of treatment effects in one regression. Figure 3 presents the results, where t-2 is the excluded time category. The figures on the left show the baseline specification, and those on the right the extended specification including trends. The four sub-figures show a similar pattern: there is a clear reform effect starting in the first implementation year (year 0), of about 0.3 years of schooling. The figures also show small positive and significant pre-reform estimates in year t-1. One possible explanation to this is the ambiguity in terms of documented starting year, where some municipalities phased in the reform over two years and as such a fraction of students were actually treated in the year before the first indicated start year. Moreover, in relation to the baseline year, treated areas have negative estimates in t-4, i.e., treated areas had a slower growth in education compared to non-treated areas four years before the reform was implemented. When controlling for implementation-year specific trends, these negative estimates move closer to zero. It is also worthwhile pointing out that in the event-study analysis, coefficients estimated a few years before (after) first implementation year are identified using subsamples of late (early) implementers. In terms of the parallel trends assumption, using the specification with trends seems to be preferable, and acknowledging that the t-1 effect is expected the diff-in-diff specification performs well. In fact, many studies have chosen to drop t-1 from the analysis to arrive at a sharper pre and post distinction. I follow this procedure in the remainder of the paper.

Figure 3. Event-study analysis of reform exposure on years of schooling



Note: Event-study estimates of reform exposure interacted with time since first treated cohort. t-2 is the left-out time category. Spikes show 95 percent confidence intervals, clustered at the municipality level.

6.4 Reform effects on years of schooling and post-compulsory schooling

The vast majority of studies of effects of the compulsory schooling reform take an interest in secondary outcomes beyond education, that come as a result of being exposed to the reform. Effects on e.g. income, health or crime can operate through longer education, but other direct channels are also possible. Quality of education, peer composition, tracking and changes to the curriculum could directly affect future outcomes. The effect of treatment, i.e. reform exposure, on length of education however serves as a proxy for a “first stage” in terms of schooling – it tells us some-

thing about how large the intervention is, for whom it bites the most, and helps to generate hypotheses regarding externalities on other outcomes. One key question also relates to whether the reform had effects beyond the new compulsory minimum.

Table 3, panel A, presents baseline estimates of the reform on years of education. In this table, results are based on coding 1, and the last pre-reform cohort has been excluded from the regressions. Column 1 shows the estimate excluding trends and indicates that the reform increased years of schooling by 0.28 years. Column 2 including trends gives an estimate of 0.31 years, or 12 percent of the standard deviation of years of schooling. Having established that the reform implied an increase in years of education, it is worth taking a closer look at the dynamics of this effect. Did the reform have differential impacts depending on socioeconomic background? Did it solely add two years of education for those at the bottom of the distribution, or did it also induce a shift beyond the new compulsory minimum? Such a spill-over effect could occur for example if the pre-reform early selection was unfavorable to talented children from disadvantaged backgrounds, which in the new school system possibly got the chance to move on further in the education system.

Columns 3 and 4 show effects on years of schooling for children with high and low educated fathers, respectively. As expected, the effect is much larger (about three times as large) among children with low educated fathers. Most students with highly educated fathers already attended education beyond compulsory level before the reform, implying that the reform had less bite in this group.

Table 3. Reform effects on educational outcomes

	(1)	(2)	(3)	(4)
	<i>All</i>	<i>All</i>	<i>High edu father</i>	<i>Low edu father</i>
A. Years of schooling				
Reform exposure coding 1	0.280*** (0.026)	0.309*** (0.016)	0.115*** (0.020)	0.361*** (0.019)
R-squared	0.013	0.014	0.002	0.021
Control group outcome mean	11.09	11.09	12.43	10.69
B. 2 year upper secondary or more				
Reform exposure coding 1	0.013*** (0.003)	0.018*** (0.002)	0.009*** (0.003)	0.018*** (0.003)
R-squared	0.014	0.014	0.004	0.018
Control group outcome mean	0.708	0.708	0.867	0.665
Number of municipalities	1,020	1,020	1,020	1,020
Muni f.e.	Yes	Yes	Yes	Yes
Cohort f.e.	Yes	Yes	Yes	Yes
Implement-spec trends	No	Yes	Yes	Yes
Observations	939,130	939,130	219,104	485,595

Note: Robust standard errors in parentheses clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Panel B shows estimates of the reform on a dummy indicating completion of two-year upper-secondary school or more.⁷ The results indicate that a small spill-over effect is present: a 0.013–0.018 higher probability to complete two years of upper-secondary education or more. As a point of comparison, the overall probability to attend two-year secondary school or any education beyond that is 0.71. Although a small effect, this indicates that reform exposure did push some individuals to higher levels beyond the expected minimum. The spill-over effect is twice as large when comparing children with low and high educated fathers (columns 3 and 4), which indicates that the shift to a comprehensive system served one of its purposes – to increase equality of opportunity.

6.5 Measurement error bias and bounding the estimates

Reform assignment comes with measurement error – in some cases it is an approximation of the starting year of the reform. As described above, some municipalities kept parallel school systems which means that there is no possibility to find a clear-cut starting point. Hence, the coding of the reform does in some cases represent an average or the majority in a given municipality and birth cohort, which will introduce measurement error in the reform indicator. Another aspect is that even though implementation might have been extensive, there was room for single individuals to apply for an exemption. We also need to assume that pupils are in the expected grade according to their age; if grade repetition or skipping a grade was a prevalent phenomenon among the affected cohorts, this is also one source of measurement error to keep in mind. To better understand the consequences of measurement error in regression analysis based on the reform, I now turn to examining the quality of reform indicators.

As a starting point to a reliability analysis of the reform coding, I acknowledge that since reform participation is a binary indicator variable, the measurement error is not classical. That is, the measurement error is correlated with the true underlying variable (Aigner 1973). The formula describing attenuation bias in the case of classical measurement error must now be modified to represent the case of non-classical measurement error.

We would like to estimate reform effects on an outcome y in the following way:

$$(2) \quad y = \alpha + \beta r^* + \varepsilon$$

Where r_{icm}^* denotes the true (unobserved) reform status of an individual i , belonging to cohort c , going to school in municipality m . In the data we observe two measures of the reform measured with error (omitting the subscripts for simplicity and following the notation in Kane et al. 1999):

$$(4) \quad E(r_1 | r^*, r_2, y) = \pi_{10} + \pi_{11} r^* \quad \text{Coding 1}$$

$$(4) \quad E(r_2 | r^*, r_1, y) = \pi_{20} + \pi_{21} r^* \quad \text{Coding 2}$$

Given the true reform participation r^* , I assume that the observed variables r_1 and r_2 are independent of each other and of y . In order for the measurement error to be classical, the further assumptions $\pi_{11} = \pi_{21} = 1$, and $\pi_{10} = \pi_{20} = 0$ must be satisfied. With a binary indicator variable these assumptions do not hold and the measurement error is correlated with the true underlying variable. We have that $\pi_{11} < 1, \pi_{21} < 1$ and $\pi_{10} > 0, \pi_{20} > 0$.

Following Aigner (1973) and Kane et al. (1999) the probability limit of β in the case of measurement error in a binary variable can be derived as follows:

$$(5) \quad p \lim \beta_{OLS} = \beta [1 - P(r^* = 1 | r = 0) - P(r^* = 0 | r = 1)]$$

⁷ Two-year secondary school is the lowest post-compulsory degree and refers to vocational post-compulsory education.

Just as in the case of classical measurement error, the OLS estimate is biased towards zero and the estimated effect is attenuated. In the case of classical measurement error, a standard remedy to inconsistencies in OLS parameters has been to use an instrumental variables strategy. With two independent measures of the variable of interest, two-stages-least squares when one measure is used as an instrument for the other produces consistent coefficients. When measurement error is non-classical, however, an IV strategy is not likely to produce consistent estimates. Nevertheless, an IV estimate can be informative, since it turns out that with non-classical measurement error it will be upward biased (Kane et al. 1999):

$$(6) \quad p \lim \beta_{2SLS} = \beta \frac{1}{\pi_{11}}$$

where r_1 has been instrumented in the first stage using r_2 as an instrument. We see that only in the case of classical measurement error ($\pi_{11} = 1$) 2SLS produces consistent estimates, and with measurement error in the categorical variable ($\pi_{11} < 1$) β_{2SLS} will be upward biased.

Thus, with measurement error in the binary indicator variable for reform participation, it turns out that both the OLS and the IV estimate (using two reform codings and instrumenting one with the other) are inconsistent, one downwards and the other upwards. Therefore, the two estimates provide a lower and an upper bound for the true parameter, and we are able to narrow down the range of possible true effects. Taking the (downward biased) estimates presented in Table 3 as a benchmark, Table 4 presents the corresponding IV estimates, where coding 2 has been used as an instrument for coding 1. Columns 1 and 3 present the OLS specifications including trends, and columns 2 and 4 the corresponding IV-2SLS estimates. The differences between the OLS and the IV estimates are relatively large; the latter are about twice as large as the OLS estimates. The take-away from this analysis is therefore that the true parameter estimate of the effect of the reform on years of schooling is within the range 0.32–0.58 years. The effect on post-compulsory education can be bounded to an increase in the range 0.021–0.046 percentage points.

Table 4. OLS and IV estimates bounding the effect of the school reform on educational outcomes

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
	Years of schooling		2 year upper secondary or more	
Reform exposure coding 1	0.316*** (0.023)	0.580*** (0.082)	0.021*** (0.004)	0.046*** (0.010)
Observations	368,480	368,480	368,480	368,480
R-squared	0.014		0.014	
Number municipalities	983	983	983	983
Muni f.e.	Yes	Yes	Yes	Yes
Cohort f.e.	Yes	Yes	Yes	Yes
Implement-spec trends	Yes	Yes	Yes	Yes
Control group outcome mean	11.09	11.09	0.708	0.708

Note: Robust standard errors in parentheses clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. In columns 2 and 4, coding 1 is instrumented with coding 2. The sample size is reduced since coding 2 is defined for a subset of the population (here 40 percent).

6.6 Can we trust staggered differences-in-differences? A robustness analysis

A number of recent papers highlight problems with a pooled difference-in-difference estimator when units are treated at different points in time, and additionally when all units are treated at some point. de Chaisemartin and D'Hautfoeuille (2020) and Goodman-Bacon (2018) show that linear regression with time- and group-fixed effects estimate weighted sums of average treatment effects in each group and time period. Treatment dummies assume that the treatment effect is constant while event-time treatment effect heterogeneity is a possibility. Borusyak and Jaravel (2018) show that short-run effects are overweighed, and long-run effects are negatively weighted. Since weights are not equal across all LATES that feed into the average, treatment effect heterogeneity might lead to a biased average effect. de Chaisemartin and D'Hautfoeuille show that weights might even be negative and that the estimate therefore can be of the wrong sign. Negative weights appear because the phase-in of treatment implies that “late” time periods can have a predicted treatment probability above 1.

To begin with, the event-study graphs in section 6.2 above show relative stability of treatment effects over time since treatment, which implies that even if long-run effects have lower weight, the pooled estimate is likely to be representative of the average effect.

Second, de Chaisemartin and D'Hautfoeuille (2020) provide a stata package to test the presence of negative weights in the context of group- and time-fixed effects models. I apply this test to the setting in this paper, i.e. the staggered difference-in-differences model estimating effects of the compulsory school reform. Using the package that checks for negative weights [twowayfeweights], I find that for all 8279 LATE:s for each combination of group and time period in my data, weights are positive. As such, the traditional difference-in-differences estimate is likely to be representative of the true average in this setting, even though treatment effect heterogeneity at the group level cannot be ruled out and might also lead to a biased average effect.

7. Summing up 15 years of research – Evidence based on the Swedish compulsory school reform

This section summarizes the papers exploiting the Swedish reform up to date. This turns out to be an impressive list of publications, published in many good journals. The papers can roughly be categorized by type of outcome; skills, income, health, crime, political decision making, financial literacy and intergenerational effects. Below I present the papers using this categorization. Some papers estimate the reduced form of the reform, while others use it as an instrument for years of education. I will return to this distinction below.

Skills. One of the basic questions in labour economics asks whether returns to education reflect higher skills and productivity that come with higher education, or if education purely works as a signal of inherent ability. Before turning to evidence on reform effects on income and other outcomes, I therefore start by presenting evidence on how the reform impacted individual skills. Lager et al. (2017) study how the reform affected men's cognitive and non-cognitive skills using data from the Swedish military enlistment. The results indicate that the reform raised cognitive skills by 5 percent of a standard deviation on average, and effects were larger among sons of fathers in low SES occupations (farmers, unqualified manual workers), and inexistant among sons of skilled workers and professionals. This result squares well with those presented in 6.4 above, which show that the reform had a larger impact on children from low SES backgrounds. The results on non-cognitive skills (emotional control) instead show negative effects on average (of 3 percent of a standard deviation), and when splitting by family background it becomes clear that the effect should be attributed to children from higher social classes. One possible interpretation of this result is that high SES sons fared worse in terms of non-cognitive skills in the comprehensive system, when exposed to a broader peer group. This finding is also in line with evidence from the Finnish reform, which indicates negative effects from late tracking on mental health for women from highly educated families (Böckerman et al. 2019).

The effects on skills and later-life outcomes may not only be explained by direct exposure to the reform and skills acquired by staying longer in school. As demonstrated by Koerselman (2013), curriculum tracking can have incentive effects for students already before the point of tracking: it creates incentives for students to work harder to be admitted to the desired track. Koerselman (2013) finds evidence of tracking effects in the U.K. – i.e. that students in the tracked system have higher test scores before the point of tracking compared to the non-tracked system. Using the Swedish compulsory school reform for a similar analysis however shows no incentive effects for the tracked system.

Finally, the reform might have affected skills through other mechanisms, for example through education spillovers to other individuals, or by “protecting” low-performing individuals from dropping out and giving them a second chance. Adermon (2013) studies sibling spillovers and finds that an older sibling exposed by the reform did not induce younger siblings to stay on longer in school. Fredriksson and Öckert (2013) study effects of school starting age and find that the educational achievement gap by school starting age is larger in the pre-reform early tracking system than post reform.

Income. Meghir and Palme (2005) presented the first reform estimates on earnings. They found a small average effect on earnings of 1.4 percent, but effects varied by parental background. Children with low educated fathers gained 3–7 percent, where the largest estimate refers to high-ability girls, and children with high educated fathers had non-trivial negative earnings effects. In light of the results by Lager et al. (2017) described above, and the literature documenting how important non-cognitive skills are for long-term outcomes, the negative earnings effects among advantaged children could be explained by a loss of socio-emotional skills rather than by a loss of cognitive skills.

More recently, Fischer et al. (2018) re-estimate the effects of the comprehensive school reform. While Meghir and Palme’s original sample was based on random samples of two cohorts, Fischer et al. use population data covering cohorts born 1938–1954 and as such cover almost the whole implementation period. The results confirm a small positive earnings effect of the reform on average, but do not lead to the same conclusions when it comes to differences by parental background: Fischer et al. find positive effects for both children of low- and high-skilled workers.

Health. More educated individuals have better health – but the causal relationship is debated. Motivated by the quest to establish whether there is a causal link between education and health several papers study reform effects on health-related outcomes.

Spasojevic (2010) builds on a relatively small survey data set and uses the reform as an instrument for years of education, focusing on outcomes measuring self-reported health (BMI and a health index combining information about both minor and severe conditions). Spasojevic’s analysis finds that one more year of education leads to a lower ill-health index and a higher probability to have a healthy BMI. Lager and Torssander (2012) study reform effects on mortality – they find that reform exposure reduced mortality risks after the age of 40. Meghir et al. (2018) study mortality, hospitalization and drug prescriptions, but find that the reform did not affect any of these outcomes. Finally, Fisher et al. (2019) study reform effects on mortality, self-reported bad health, and smoking and similarly find no effects on these health outcomes. A possible explanation as to why the results on health and mortality differ between studies is that the two latest studies use larger samples and a longer follow-up periods.

Finally, Palme and Simeonova (2015) study a more specific health outcome: incidence and mortality of breast cancer – a form of cancer positively associated with education and labelled a “welfare disease”. They find that the reform increased the risk of women being diagnosed with breast cancer, and lead to an elevated probability of death from breast cancer.

Crime. Individuals engaging in criminal activity are on average low educated and a key policy question is whether educational interventions can help to protect individuals from a criminal career. Education can raise the returns to legal activities, can lead to a different peer group, and also protects individuals directly through the incapacitation effect while still in education.

Hjalmarsson et al. (2015) study the reform effect through years of schooling on convictions and incarcerations. One additional year of schooling induced by the reform decreased the likelihood of conviction by 6.7 percent, and incarceration by 15.5 percent among men. Meghir et al. (2012) study reform effects on crime both in the treated generation and among their offspring – i.e., the intergenerational effect of the reform on crime. The study finds that the policy reduced crime rates both for the targeted generation and among their children. Reform exposure reduced the probability of ever being convicted by 5 percent among men born 1954–1955. Sons of exposed men were also affected and had a lower probability of ever being convicted.

Political participation. Lindgren et al. (2017) take an interest in the social divide in political participation and ask whether the comprehensive schooling reform was an effective policy to raise political involvement among individuals from low SES backgrounds. The study focuses on political candidacy: it uses information on all nominated and elected candidates in six parliamentary, county council and municipal elections in Sweden between 1991 and 2010. The reform significantly increased the probability of political candidacy among individuals from working class backgrounds – the impact of family background on the likelihood of seeking public office was reduced by up to 40 percent.

Financial decision-making. More educated individuals participate in financial markets – even after controlling for wealth or income. Black et al. (2018) ask whether there is a causal link between education and investment behavior, and whether policies that increase educational attainment also change people's investment behavior. The study uses the reform as an instrument for years of education to analyze the effect of one more year of schooling on stock market participation. The results show that one more year of education increases stock market participation among men by 2 percentage points (over a baseline of 42 percent), and a suggestive channel explaining this result is greater financial wealth. There is no evidence that reform-induced changes in education affected women's financial decision-making, nor that there were spillover effects to children.

Girishina (2019) studies the effects of education on wealth. Using the reform as an instrument for education she finds that one more year of schooling increases the value of an individual's total assets, which is consistent with Black et al.'s conclusion.

Intergenerational spillovers. Strong intergenerational correlations are observed on a range of outcomes, such as education, earnings, health and criminal involvement. These correlations can be the result of direct causal influences from parents to children, through investment behavior or role model mechanisms, but are also explained by other underlying characteristics shared by parents and children that give rise to correlations in outcomes. Because of the difficulty in distinguishing between direct causal spillovers and correlations, the compulsory school reform has offered an opportunity to learn something about intergenerational spillovers of exposure to comprehensive schooling on offspring's outcomes.

Holmlund et al. (2011) study the effects of parents' schooling on children's schooling. The intergenerational correlations, regressing years of schooling of children on years of schooling of fathers/mothers, yields estimates at 0.23 and 0.28 years for fathers and mothers, respectively. Using the reform as an instrument for education, they find that one more year of schooling among parents leads to about 0.1 more years in the next generation, which is much lower than the correlations in the data. Also Sikhova (2019) shows that the reform affected intergenerational transmissions of human capital.

Lundborg et al. (2014) study intergenerational effects on sons' cognitive and non-cognitive outcomes, as measured at military enlistment at the age of 18. The study uses an IV approach and finds positive effects of maternal education on sons' cognitive ability and a global health index, but no effects of father's education. The effect magnitudes indicate that one more year of maternal schooling increases both cognitive ability and global health by about 0.1 of a standard deviation.

Lundborg and Majlesi (2018) consider that intergenerational transmissions also can run from child to parent. Using the reform as an instrument for education, they study how education in the treated generation affects health outcomes among elderly parents. If more education provides children with better resources to care for their elderly parents, it is possible that parental health will be affected. On average, the reform-induced changes in years of schooling do not affect parents'

longevity, but there is some heterogeneity suggesting that low educated fathers are positively affected by their offspring's education – they exhibit higher survival at age 70–80 when their children are exposed to longer compulsory education.

Nybom and Stuhler (2016) study trends in intergenerational mobility and ask how such trends should be interpreted: changes in mobility can be the result of current policies, but can also be explained by events in the past that have dynamic effects on mobility over several generations. As an example, a shift towards a more meritocratic society increases mobility when highly skilled children from poor families are rewarded for their skills rather than their background. In the next generation however, it is possible that mobility declines again, since skills are passed on from parents to their children. To illustrate their structural model, Nybom and Stuhler show that individuals exposed to the Swedish compulsory school reform exhibit higher intergenerational mobility, but that the reform decreased mobility in the next generation.

7.1 IV and the exclusion restriction

The reform extended compulsory education, but also implied changes to the curriculum, affected tracking, and consequently the peer group composition. The effects of the reform on future outcomes should therefore be interpreted as the total (net) effect, incorporating all possible channels through which reform exposure affected outcomes. Despite this, a number of the papers mentioned above use the reform as an instrument for years of education (see e.g. Spasojevic 2010, Lundborg et al. 2014, Hjalmarsson et al. 2015, Black et al. 2018, Lundborg and Majlesi 2018), although the exclusion restriction is unlikely to hold. If we are primarily interested in understanding the effects of the Swedish reform on a variety of outcomes, the reduced form vs. IV distinction may not be so problematic – we can in most cases back out the reduced form, and the main question is to identify whether there is a significant effect, its sign, and which subgroups are mostly affected. Using the reform as an instrument for years of schooling is more problematic when we start interpreting the effect as caused by years only, and when comparing the estimates with those in the literature on returns to schooling. The estimates based on the Swedish reform are unlikely to be comparable to those of other studies both because the exclusion restriction is unlikely to hold, and because the compliers come from the lower tail of the education distribution. If there are heterogeneous treatment effects, it is not clear that we can compare studies using compulsory schooling reforms with those using e.g. college openings as instruments for years of schooling.

8. Conclusions

This paper demonstrates how a natural experiment can be used to estimate causal effects. The Swedish compulsory school reform was implemented gradually across cohorts and municipalities, which under certain assumptions makes it possible to estimate causal effects of the reform. Many papers have used differences-in-differences specifications to estimate effects of the reform on a range of outcomes, showing that extending basic education provided children from low socioeconomic background with better opportunities in life. Not only did they attain higher levels of education – they also earned higher earnings, were less likely to participate in crime, and more likely to run for office. But there is also some evidence in support of the notion that a comprehensive system could be harmful to high ability children. Meghir and Palme (2005) found negative earnings effects for high SES groups, although this result is not replicated by Fisher et al. (2018). Evidence both from Sweden (Lager et al. 2017) and Finland (Böckerman et al. 2019) indicate negative effects of comprehensive schooling on non-cognitive skills for children from higher socioeconomic backgrounds. The reform had winners and losers, but it undoubtedly reduced inequality.

Earlier academic work has argued that the reform was a failure, on the account that recruitment to tertiary education dropped among youth with working-class backgrounds for cohorts affected by the reform (Rothstein 1996).⁸ But this conclusion does not account for the possibility that the counterfactual outcome could have been even less encouraging

⁸ Earlier work with a different conclusion regarding the reform outcome is Erikson (1996) who found that the introduction of the comprehensive school coincided with reduced social inequality in education.

from an equality point of view and highlights the importance of rigorous evaluation methods. As I have shown in the empirical part of this paper, the reform increased the probability that children stayed in school beyond compulsory level, and this effect is twice as large for children from low educated families compared to high educated families. The reform also reduced the gap in cognitive skills between children of different social background. In this light, the reform must be considered successful in reaching its objective of reducing social inequalities. □

References

- Acemoglu, Daron and Joshua D. Angrist (2000), “How Large are Human-Capital Externalities? Evidence from Compulsory Schooling Laws”, in Bernake, Ben and Kenneth Rogoff (eds.), NBER Macroeconomics Annals 2000, Volume 15.
- Adermon, Adrian (2013), “Sibling Spillovers in Education”, in “Essays on the Transmission of Human Capital and the Impact of Technological Change”, Doctoral thesis, Uppsala university.
- Aigner, Dennis J. (1973), “Regression with a Binary Independent Variable Subject to Errors of Observation”, *Journal of Econometrics* 1: 49–60.
- Angrist, Joshua D. and Alan B. Krueger (1991), “Does Compulsory School Attendance Affect Schooling and Earnings?”, *Quarterly Journal of Economics* CVI(4): 979–1014.
- Arendt, Jacob Nielsen (2005), “Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification”, *Economics of Education Review*, 24(2), pp. 149–160.
- Ashenfelter, Orley and Alan B. Krueger (1994), “Estimates of the Returns to Schooling from a New Sample of Twins”, *American Economic Review* 84(5): 1157–1173.
- Behrman, Jere R. and Mark R. Rosenzweig (1992), “Does Increasing Women’s Schooling Raise the Schooling of the Next Generation?”, *American Economic Review* 92(1): 323–334.
- Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes (2005), “Why the Apple Doesn’t Fall Far: Understanding Intergenerational Transmission of Human Capital”, *American Economic Review* 95(1): 437–449.
- Black, Sandra E., Paul J. Devereux, Petter Lundborg and Kaveh Majlesi (2018), “Learning to Take Risks? The Effect of Education on Risk-Taking in Financial Markets”, *Review of Finance*, 22(3), pp. 951–975.
- Borusyak, Kirill and Xavier Jaravel (2018), “Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume”, unpublished manuscript.
- Bound, John and David A. Jaeger (1996), “On the Validity of Season of Birth as an Instrument in Wage Equations: A Comment on Angrist & Krueger’s “Does Compulsory School Attendance Affect Schooling and Earnings?”, NBER working paper 5835.
- Böckerman, Petri, Mika Haapanen, Christopher Jepsen and Alexandra Roulet (2019), “School Tracking and Mental Health”, IZA Discussion paper No. 12733.
- Chevalier, Arnaud (2004), “Parental Education and Child’s Education: A Natural Experiment”, IZA Discussion Paper No. 1153.
- Currie, Janet and Enrico Moretti (2003), “Mother’s Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings”, *The Quarterly Journal of Economics* 118(4): 1495–1532.
- de Chaisemartin, Clément and Xavier D’Hautfoeuille (2020), “Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects”, forthcoming, *American Economic Review*.
- Educational Bureau (1960-1964), Tables of pupils in the compulsory school.
- Erikson, Robert and Rune Åberg, eds. (1984), *Välfärd i förändring. Levnadsvillkor i Sverige 1968-1981*. Stockholm. (English translation: *Welfare in Transition*. Oxford: Clarendon Press, 1987.)
- Erikson, Robert and Jan O. Jonsson (1996), “Introduction. Explaining Class Inequality in Education: The Swedish Test Case”, in Erikson, Robert. and Jan. O. Jonsson (eds.), *Can Education Be Equalized?*, Westview Press.
- Erikson, R. (1996), “Explaining Change in Educational Inequality – Economic Security and School Reforms”, in Erikson, Robert and Jan O. Jonsson (eds.), *Can Education Be Equalized?*, Westview Press.

- Fischer, Martin, Gawain Heckley, Martin Karlsson and Therese Nilsson (2018), "Did Sweden's Comprehensive School reform reduce Inequalities in Earnings?" in Fisher, Martin, *The Long-Term Effects of Education on health and Labor Market Outcomes – Evidence from Historical School Reforms in Sweden and Germany*, PhD dissertation, University of Duisburg-Essen.
- Fischer, Martin, Ulf-G Gerdtham, Gawain Heckley, Martin Karlsson, Gustav Kjellsson and Therese Nilsson (2019), "Education and Health: Long-run Effects of peers, Tracking and Years", IFN Working Paper No. 1300.
- Fredriksson, Peter and Björn Öckert (2014), "Life-Cycle Effects of Age at School Start", *The Economic Journal*, 124, September, pp. 911–1004.
- Girshina, Anastasia (2019), "Wealth, Savings, and Returns Over the Life-Cycle: The Role of Education", Job Market Paper, Stockholm School of Economics.
- Goodman-Bacon, Andrew (2018), "Difference-in-differences with Variation in Treatment Timing", NBER working paper 25018.
- Hjalmarsson, Randi, Helena Holmlund and Matthew J. Lindquist (2015), "The Effect of Education on Criminal Convictions and Incarceration: Causal Evidence from Micro-data", *The Economic Journal*, 125 (September), pp. 1290–1326.
- Holmlund, Helena (2007), "A Researcher's Guide to the Swedish Compulsory School Reform", SOFI working paper 9/2007.
- Holmlund, Helena, Mikael Lindahl and Erik Plug (2011), "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods", *Journal of Economic Literature*, 49(3), pp. 615–651.
- Härnqvist, Kjell (2000), "Evaluation through Follow-up. A Longitudinal Program for Studying Education and Career Development", in C.-G. Janson, ed., *Seven Swedish longitudinal studies in behavioral science*. Stockholm, Forskningsrådsnämnden.
- Kane, Thomas J., Cecilia Elena Rouse and Douglas Staiger (1999), "Estimating Returns to Schooling When Schooling is Misreported", Working Paper 419, Industrial Relations Section, Princeton University.
- Koerselman, Kristian (2013), "Incentives from Curriculum Tracking", *Economics of Education Review*, 32, pp. 140–150.
- Lager, Anton and Jenny Torssander (2013), "Causal Effect of Education on Mortality in a Quasi-Experiment on 1.2 million Swedes", *PNAS*, 109(22), pp. 8461–8466.
- Lager, Anton, Dominika Seblova, Daniel Falkstedt and Martin Lövdén (2017), "Cognitive and emotional outcomes after prolonged education: a quasi-experiment on 320 182 Swedish boys", *International Journal of Epidemiology*, 46(1), pp. 303–311.
- Lindgren, Karl-Oskar, Sven Oskarsson and Christopher T. Dawes (2017), "Can Political Inequalities Be Educated Away? Evidence from a Large-Scale Reform", *American Journal of Political Science*, 61(1), pp. 222–236.
- Lleras-Muney, Adriana (2005), "The Relationship between Education and Adult Mortality in the U.S.", *Review of Economic Studies* 72(1): 189–221.
- Lochner, Lance and Enrico Moretti (2004), "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests and Self-Reports", *American Economic Review* 94(1): 155–189.
- Lundborg, Petter, Anton Nilsson and Dan-Olof Rooth (2014), "Parental Education and Offspring Outcome: Evidence from the Swedish Compulsory School Reform", *American Economic Journal: Applied Economics*, 6(1), pp. 253–278.
- Lundborg, Petter and Kaveh Majlesi (2018), "Intergenerational Transmission of Human Capital: Is it a One-Way Street?", *Journal of Health Economics*, 57, pp. 206–220.
- Marklund, Sixten (1980), "Från reform till reform: Skolsverige 1950-1975, Del 1, 1950 års reformbeslut", Skolöverstyrelsen och Liber Utbildnings Förlaget.
- Marklund, Sixten (1981), "Från reform till reform: Skolsverige 1950-1975, Del 2, Försöksverksamheten", Skolöverstyrelsen och Liber Utbildnings Förlaget.
- Maurin, Eric and Sandra McNally (2005), "Vive La Révolution! Long Term Returns of 1968 to the Angry Students", IZA Discussion Paper No. 1504.
- Meghir, Costas and Mårten Palme (2005), "Educational Reform, Ability and Family Background", *American Economic Review* 95(1): 414–424.
- Meghir, Costas, Mårten Palme and Marieke Schnabel (2012), "The Effect of Education Policy on Crime: An Intergenerational Perspective", NBER Working paper 18145.
- Meghir, Costas, Palme, Mårten and Emilia Simeonova (2018), "Education and Mortality; Evidence from a Social Experiment", *American Economic Journal: Applied Economics*, 10(2), pp. 234–256.

- National Board of Education (Skolöverstyrelsen) (1954–1962), “Redogörelse för försöksverksamhet med enhetsskola”, Aktuellt från Skolöverstyrelsen.
- National Board of Education (Skolöverstyrelsen) (1960), “Enhetsskolan under 10 år, Kort redogörelse för försöksverksamheten läsåren 1949/50 – 1958/59”, Kungliga skolöverstyrelsens skriftserie 46.
- Nybom, Martin and Jan Stuhler (2016), “Interpreting Trends in Intergenerational Mobility”, manuscript. Previous version published as SOFI Working paper 3/2014.
- Oreopoulos, Philip, Marianne E. Page and Ann Huff Stevens (2006), “The Intergenerational Effects of Compulsory Schooling”, *Journal of Labor Economics* 24(4): 729–760.
- Palme, Mårten and Emilia Simeonova (2015), “Does Women’s Education Affect Breast Cancer Risk and Survival? Evidence from a population Based Social Experiment in Education”, *Journal of Health Economics*, 42, pp. 115–124.
- Pekkarinen, Tuomas, Roope Uusitalo and Sari Pekkala (2006), “Education Policy and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform”, IZA Discussion Paper No. 2204.
- Rothstein, Bo. (1998), “The Social Democratic State: Swedish Model and the Bureaucratic Problem of Social Reforms”, University of Pittsburgh Press.
- Sikhova, Aiday (2019), “Better Parents or Richer Parents: understanding Intergenerational Transmission of Human Capital”, manuscript.
- Skatteverket (2007), “Sveriges församlingar genom tiderna”, <http://www.skatteverket.se/folkbokforing/sverigesforsamlingargenomtiderna.4.18e1b10334ebe8bc80003817.html>.
- Spasojevic, Jasmina (2010), “Effects of Education on Adult Health in Sweden: Results from a Natural Experiment”, *Contributions to Economic Analysis*, vol 290, Chapter 9.
- Statistics Sweden (1968–1969), Tables of pupils in the compulsory school, Statistiska meddelanden U1968:2, U1969:5.
- Viarengo, Martina (2007), “An Historical Analysis of the Expansion of Compulsory Schooling in Europe after the Second World War”, Working paper 97/07, Department of Economic History, London School of Economics.

APPENDIX

A.1 Reform coding for register data

There are two independent ways to obtain a reform code to attach to register data, one is to use available documentation on when the reform was in place, the other is to deduce it from a register-based data set of large sample size.

A.1.1 Reform coding based on documentation [CODING 1]

To obtain a complete code of the implementation of the nine-year comprehensive school, I have, with help from Anders Björklund, Valter Hultén and Mikael Lindahl, used two main sources:

a. Marklund and the National Board of Education

Marklund (1981) provides a list with the quantitative development of the reform from 1949/50 to 1960/61. This documentation states which municipalities (or which parts of a municipality) that entered the assessment programme each year. However, Marklund (1981) does not list which grades that were exposed to the reform, in each municipality and each year. This information is available in the yearly publications that the National Board of Education published during the course of the trial period (Aktuellt från Skolöverstyrelsen 1953-1962). These publications summarized many of the aspects of the ongoing educational development, one of which was the participating municipalities and also which grades that were subject to the reform. These publications together cover the years 1951/52 to 1960/61. It is noteworthy that these two sources, Marklund and the Board of Education reports, in general coincide in terms of municipalities listed. There is one difference in that the yearly publications from the Board of Education list a few more municipalities as participating in the reform than what is mentioned in Marklund.

In the guidelines for the reform assessment, it was stated that only pupils in grades 1 through 5 would be subject to any changes. Therefore the above information from Marklund and the Board of Education, that covers almost the whole assessment period with the last year being 1960/61, makes it possible to assign whether individuals born in 1938 to 1949 were subject to the reform or not (the 1938 cohort was the first one to be affected, 1949 is the cohort of 5th graders in 1960/61). From 1961/62 these sources do not tell us what is going on, but we know for sure that pupils in 6th grade and above should not be subject to any changes. Therefore, we can code the cohort of 6th graders in 1961, and older pupils (the 1949 cohort and older), but for younger cohorts there could be changes from 1961 onwards that are not captured by these sources.

b. The Educational Bureau (1960–1964) and Statistics Sweden (1968–1969)

When the final decision about the complete introduction of the new school was taken, in 1962, the experimental period also came to an end. Now, municipalities were required to implement the reform, but a transition period allowed them to postpone the implementation, however no longer than until 1969. Thus, also in the early 1960s there is some variation in reform implementation, affecting cohorts born from 1950 onwards. Marklund (1981) and the publications from the Board of Education were mainly concerned with the assessment programme, and thus they do not document reform implementation in the 1960s. To use the variation in compulsory schooling legislation for the 1950s cohorts, it is possible to trace reform implementation in the early 1960s from municipality tables from Statistics Sweden (1968, 1969) and from the Educational Bureau (Undervisningsbyrån) (1960, 1961/62, 1963/64). For each municipality, the tables from the Educational Bureau give the number of pupils in each grade in both the old school (folkskolan) and the new nine-year comprehensive school (grundskolan). From such a table it is possible to see in which grade and year the implementation took place. See the examples following below.

Example 1

Municipality m in year t, cohort size is around 500.

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	500	500	500	0	500	500	0	0
Grundskola (new)	500	0	0	0	500	0	0	0	0

Municipality m in year t+1

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	0	500	500	0	0	500	0	0
Grundskola (new)	500	500	0	0	500	500	0	0	0

Municipality m in year t+2

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	0	0	500	0	0	0	0	0
Grundskola (new)	500	500	500	0	500	500	500	0	0

Municipality m in year t+3

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	0	0	0	0	0	0	0	0
Grundskola (new)	500	500	500	500	500	500	500	500	0

From these tables it is possible to conclude that the cohort of 5th graders in year t, that is, the cohort born in t-11, is the first cohort in municipality m, to be affected by the reform. All younger birth cohorts were also affected (since even if you were in grades 2–4 in year t, you would eventually reach grade 5 and thus be phased into the new school).

Example 1 is a stylized example, in reality the tables year-by-year might look either as in example 2 or 3 below.

Example 2

Municipality m in year t

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	500	500	500	250	500	500	0	0
Grundskola (new)	500	0	0	0	250	0	0	0	0

Municipality m in year t+1

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	0	500	500	0	250	500	0	0
Grundskola (new)	500	500	0	0	500	250	0	0	0

In this case, it is not clear which cohort that should be assigned as the first reform cohort. Is it the cohort in 5th grade in year t or in t+1? In these cases the reform implementation has been set to start when at least half of a cohort is facing the reform. However, it is clear that the coding here will introduce some measurement error.

Example 3

Municipality m in year t

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	500	500	500	0	500	500	0	0
Grundskola (new)	500	0	0	0	500	0	0	0	0

Municipality m in year t+1

Grade	1	2	3	4	5	6	7	8	9
Type of school	Nr of pupils								
Folkskola (old)	0	0	500	500	0	0	0	0	0
Grundskola (new)	500	500	0	0	500	500	500	0	0

This example shows that the tables are not always coherent between years. In year t, it looks like the first cohort is the fifth graders in t, whereas in year t+1, it seems like the first cohort is the one of 7th graders in t+1. In these cases, the information on which cohort entered first is taken from the last table that reveals a shift between the old and the new school (in the light of the example above, it would be the 7th graders in year t+1, that is the cohort born in t+1-13).

Note that the first table from the Educational Bureau is from 1960/61, which means that municipalities that introduced the reform very early cannot be coded using this second source of information. That is, in the case all pupils in grade 1

through 9 were already in the new school in 1960, it is not possible to see when the shift took place. In those cases, we rely solely on the first source (Marklund). Luckily, there is some overlap between the two sources: for 158 municipalities I have obtained a coding from both Marklund and the Bureau. In 9 out of 158 cases, the coding differs between these sources, and in those cases, I have used Marklund.

In some cases we know from Marklund that a municipality introduced the reform in different parts of the municipality at different points in time. If these were early implementers, the statistical tables do not reveal when the majority of the pupils in a municipality were shifted into the new system. In that case, Marklund states which school district, or which schools within a municipality that introduced the new school, it is possible to assign these schools to a sub-region of the municipality (a parish). There are however, a few municipalities where we know that the reform was introduced gradually, but there is no information on which schools or which part of the municipality. These municipalities cannot be coded and must be dropped from the sample: Hälsingborg, Jönköping, Linköping, Skellefteå, Sundbyberg and Södertälje.

The three big cities, Stockholm, Göteborg and Malmö, are also problematic to code. They implemented the reform at different points in time in different parts of their municipalities, and the coding has been constructed as follows (note that information on parish is necessary)⁹:

Stockholm. From the statistical tables from the Educational Bureau, it is clear that in 1962, the whole cohort of 8th graders (the 1948 cohort) was shifted into the new comprehensive school. However, reform implementation had started gradually earlier, at first in the southern suburbs of Stockholm. Based on information on parish of residence, the south suburbs can be dropped, and the change that affects (approximately all the rest) of Stockholm can be coded and the first cohort affected is set to 1948.

Göteborg. The first cohort where all pupils are in the new school is the 1950 cohort. Early implementing parishes are dropped.

Malmö. The first cohort where all pupils are in the new school is the 1949 cohort. Early implementing parishes are dropped.

The procedure outlined above allows me to find the first cohort affected by the reform in almost all municipalities. Some could not be coded due to ambiguity as to which part of the municipality implemented the reform (mentioned above). Yet another three municipalities could not be coded, simply because they did not show up in the statistical records: Fjälkinge, Svarteborg and Sörbygden.

A.1.2 Reform coding deducted from large-sample register data [CODING 2]

With a large enough sample, it is possible to adopt the following strategy to find out when a municipality implemented the reform: split the sample (randomly) in two parts (in this paper 60/40 percent, sampling by municipality to maintain proportions at the group level). I use the 60 percent sample to identify the reform date, and 40 percent for estimation. Within the 60 percent sample, I drop all individuals with education higher than the new compulsory minimum (grundskola), using the information on completed education levels in Statistics Sweden's education register. Now we are left with only observations of the old minimum (folkskola) and the new minimum (grundskola). Assign a dummy equal to one for the new comprehensive school. Collapse this data by birth cohort and municipality, and look at the average of the comprehensive school dummy for each cell. With a clean-cut implementation, we should observe that within a municipality, the average shifts from 0 to 1 between two specific cohorts, and this is when the reform is implemented. In reality, the cohort-to-cohort changes are not always so clean, and one can assign a reform to cohorts where the cohort/municipality-average of the compulsory school dummy is ≥ 0.5 .

⁹ Assigning the reform based on information of parish of residence is an approximation, where I map a given school (or school district) to the parish/es in which it lies.

With this procedure it might be the case that you assign the reform to cohort t , but in cohort $t+1$ the dummy average is <0.5 and for cohort $t+2$ it shifts back to ≥ 0.5 . In the empirical part of this paper I have, in the case I observe more than one shift, assigned the reform to the shift that never moves back below 0.5.

Note also that some municipalities did not implement the reform uniformly within itself; most notably this was the case in the big cities. To arrive at a cleaner definition I therefore drop “early implementing” parishes in the big cities, and other unclear parishes/municipalities where it is known that implementation was gradual (see coding 1 above).