

DISCUSSION PAPER SERIES

No. 8646

**THE EFFECT OF EDUCATION ON
CRIMINAL CONVICTIONS AND
INCARCERATION: CAUSAL
EVIDENCE FROM MICRO-DATA**

Randi Hjalmarsson, Helena Holmlund
and Matthew Lindquist

PUBLIC POLICY



Centre for Economic Policy Research

www.cepr.org

Available online at:

www.cepr.org/pubs/dps/DP8646.asp

THE EFFECT OF EDUCATION ON CRIMINAL CONVICTIONS AND INCARCERATION: CAUSAL EVIDENCE FROM MICRO-DATA

Randi Hjalmarsson, Queen Mary, University of London and CEPR
Helena Holmlund, Stockholm University
Matthew Lindquist, Stockholm University

Discussion Paper No. 8646
November 2011

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **PUBLIC POLICY**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Randi Hjalmarsson, Helena Holmlund and Matthew Lindquist

CEPR Discussion Paper No. 8646

November 2011

ABSTRACT

The Effect of Education on Criminal Convictions and Incarceration: Causal Evidence from Micro-data

This paper studies the causal effect of educational attainment on conviction and incarceration using Sweden's compulsory schooling reform as an instrument for years of schooling and a 25 percent random sample from Sweden's Multigenerational Register matched with more than 30 years of administrative crime records. The first stage of the analysis employs a differences-in-differences design to account for the non-random implementation of the reform across municipalities, and finds that exposure to the reform increased average educational attainment by 0.28 years for males and 0.16 years for females. Our 2SLS estimates indicate that more schooling has a significant negative effect on convictions and incarceration at both the extensive and intensive margins. These effects are generally seen for both males and females. Specifically, one additional year of schooling decreases the likelihood of incarceration by 16 percent for males and the likelihood of conviction by 7.5 and 11 percent for males and females, respectively. In addition, we find that the effect of education on crime persists across birth cohorts, throughout the life cycle, and across crime categories.

JEL Classification: I2 and K42

Keywords: crime, education and school reform

Randi Hjalmarsson
Queen Mary, University of London
School of Economics and Finance
Mile End Road
London

Helena Holmlund
SOFI
Stockholm University
SE-106 91 Stockholm
SWEDEN

Email: r.hjalmarsson@qmul.ac.uk

Email: helena.holmlund@sofi.su.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=172194

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=162914

Matthew Lindquist
SOFI
Stockholm University
SE-106 91 Stockholm
SWEDEN

Email: Matthew.Lindquist@sofi.su.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=173456

Submitted 25 October 2011

1. Introduction

There is an extensive amount of research dedicated towards identifying the private economic returns to education.¹ In contrast, much less attention has been given to studying the impact of education on other non-economic outcomes.² It is necessary to know the size of the external (social) benefits of education, however, when evaluating policy initiatives. The current paper will contribute to this literature by studying the impact of education on one such non-economic outcome – criminal activity – using register data from Sweden.

There is much evidence that criminals tend to be less educated than the rest of the population. For instance, according to a 2003 Bureau of Justice Statistics Report, about 41 percent of inmates in U.S. prisons or jails in 1997 had not completed high school (or its equivalent) in comparison to just 18 percent of the general population aged 18 or older (Harlow, 2003).³ In addition, this phenomenon is not isolated to the United States. Using Census data, Machin, Marie, and Vujić (2011) show that 2.57 percent of UK men aged 21-25 with no educational qualifications were in prison in 2001 compared to 0.3 percent of those with some qualifications. The data for the current project indicates the same pattern. Using a 25 percent random sample of Swedes born between 1943 and 1955 (more than 400,000 individuals in total), we see that the average years of schooling for males (females) with at least one conviction is 10.8 (11.4) while the average years of schooling for males (females) with no convictions is 11.5 (11.7).

¹ See Card (1999) for a review.

² One of the more studied non-pecuniary outcomes is health: Grossman and Kaestner (1997) and Lleras-Muney (2005) find a positive relationship between education and health outcomes. See Oreopoulos and Salvanes (2009) for a review of the literature; they argue that the non-pecuniary returns to education are at least as large as the pecuniary returns to education.

³ Petit and Western (2004) report that 40 percent of state prisoners in 1997 lacked a high school diploma.

But, are the above described relationships causal in nature? Theoretically, there are a number of reasons to expect that an increase in education causes a decrease in crime. Two recent reviews of the economics of education and crime literature by Lochner (2008, 2010) highlight these potential underlying mechanisms.^{4,5}

First, education increases wages and, therefore, the opportunity costs to committing a crime. Individuals cannot be engaged in the legitimate labor market during the time allocated towards the planning and execution of their crime or when detained at the police station upon arrest, jailed prior to conviction, or incarcerated post-conviction. Second, as highlighted by Lochner and Moretti (2004) and Lochner (2010), individuals may learn to be more patient through schooling. Such individuals place more weight on their potential future earnings and, thus, may be more likely to factor in the chances of getting caught and the expected prison sentence when deciding whether to commit a crime. Becker's economic model of crime (1968) implies that the increased opportunity costs and patience associated with higher education would decrease criminal activity. Third, increased schooling can decrease the chance that an individual engages in criminal activity by increasing his attachment to legitimate society. In addition, the more educated an individual is, the more educated his peers are likely to be. If education decreases crime through any of the above mechanisms, then there may also be a social multiplier or peer effect that further decreases criminal activity.⁶

⁴ Lochner (2010) formalizes much of this discussion with a simple model of the choice individuals face between legitimate work and criminal activity. Education in this model is a human capital investment that increases future legitimate work opportunities and discourages participation in crime.

⁵ This study focuses on the effect of education on an individual's future participation in crime. However, there can also be a *contemporaneous* effect as youths are 'incapacitated' while sitting in school. See Jacob and Lefgren (2003) and Luallen (2006) for empirical studies of this question.

⁶ It is not impossible, however, for education to increase crime, though this may be limited to certain types of criminal activity such as white collar crime. In this case, schooling may provide individuals with the skills necessary to commit these crimes as well as the job opportunities in which such white collar crimes could be committed. Lochner's (2004) human capital model of crime implies that the relationship between education and white collar crime should at least be less negative than that between 'unskilled' crimes, and may even be positive.

A number of empirical issues, however, make it difficult to interpret the above described education-crime relationships as causal. Perhaps the most important source of endogeneity is unobservable heterogeneity. A causal interpretation can be complicated by the possibility that the observation of a negative correlation arises because of unobserved individual characteristics, such as low risk aversion, lack of patience, or low ability, that simultaneously place individuals at high risk of *both* crime and low educational outcomes. Another potential source of endogeneity is reverse causality. How much of the observed negative correlation is driven by the effect of an individual's criminality on his education outcomes? Hjalmarsson (2008) provides some evidence that this is a valid concern, at least with regards to crimes committed during the teenage years.⁷

Lochner and Moretti (2004) were the first to address these issues and provide convincing empirical estimates of the causal effect of education on crime. They study the causal effects of schooling on incarceration and arrests by instrumenting for educational attainment (years of education or high school graduation) with compulsory schooling laws in the state of birth and year when the individual was 14 years old. They find that one extra year of schooling results in 0.10 and 0.37 percentage point reductions in the chance of incarceration for whites and blacks, respectively. They also find that a one year increase in average schooling reduces both property and violent crimes by 11 to 12 percent.⁸ More recently, Machin, Marie and Vujić (2011) used changes in the minimum school leaving age in Britain from 14 to 15 in 1947 and 15 to 16 in

⁷ Another empirical issue is that many papers use administrative arrest and incarceration records. These measures are only observed, however, for individuals who are caught. To the extent that more educated individuals are 'better' criminals, i.e. they are less likely to be caught, convicted, and/or incarcerated conditional on arrest, this can bias estimates of the relationship between education and administrative crime measures.

⁸ Lochner (2004) uses a similar approach to look at the effect of schooling on arrest rates for white collar crime and finds a positive, though insignificant, relationship.

1973 as an exogenous source of variation in the years of schooling. They find a significant, negative effect of education on property crimes and an insignificant effect on violent crimes.⁹

In a recent working paper – perhaps the first to use micro-data to study the causal impact of education on crime – Meghir, Palme and Schnabel (2011) study the impact of Sweden’s compulsory schooling reform on crime. The use of this reform as a source of exogenous variation in education was pioneered by Meghir and Palme in their 2005 study, which looked at the effect of the reform on educational attainment and earnings.¹⁰ The Swedish compulsory school reform, which primarily extended compulsory schooling from seven to nine years, differs from the U.S. and U.K. reforms studied by Lochner and Moretti (2004) and Machin, Marie and Vujić (2011), respectively. The Swedish reform was implemented at different times across municipalities during the 1950s and 60s, which allows Meghir, Palme and Schnabel (2011) to estimate the effect of the school reform net of any general equilibrium effects that the reform may have had on the Swedish labor market. In their experiment, they compare individuals from the same birth cohort who are working in the same labor market, but who happened to have been exposed to two different school systems.

Meghir, Palme and Schnabel (2011) study the effect of this reform on the crime of males directly affected by the reform and on the sons of those affected by the reform. This intergenerational perspective is, perhaps, the key distinguishing factor that sets their work apart from the previous studies by Lochner and Moretti (2004) and Machin, Marie and Vujić (2011). Using a reduced form specification, they find evidence of a negative effect of the reform on both the likelihood of conviction (a 5 percent reduction) and the number of convictions (a reduction of

⁹ Buonanno and Leonida (2009) find that a 10 percentage point increase in high school graduation rates would reduce property crime rates by 4% in a study using panel data for 20 Italian regions from 1980 to 1995. Though they do not have an exogenous source of variation in education, they control for region and time fixed effects, region specific quadratic time trends, and a vector of time-varying region specific characteristics.

¹⁰ Earlier versions of this paper include Meghir and Palme (1999) and Meghir and Palme (2001).

0.25 crimes for those coming from low SES backgrounds), but not on imprisonment. Perhaps more striking is their result that sons whose fathers were assigned to the school reform have a 2.5 percent lower probability of being convicted. They argue that these intergenerational effects operate through improved parenting and investments in children.

In the current paper, we also study the effect of educational attainment on crime using Sweden's compulsory schooling reform. Our analysis uses a 25 percent random sample of those born between 1943 and 1955 from Sweden's Multigenerational Register (more than 400,000 individuals). Longitudinal data concerning income, education, parish and municipality of residence in 1960, as well as criminal convictions and sentences for each individual for the years 1973 to 2007 were matched on to the data.

Because municipalities were not randomly selected into the school reform experiment and evaluation program, it is likely that the reform is correlated with municipality specific characteristics. Thus, our analysis is based on a differences-in-differences design, which includes: (i) birth cohort fixed effects, (ii) municipality fixed effects to control for municipality specific characteristics that are constant over time, and (iii) municipality specific time trends to control for municipality characteristics that change over time. Though it is still possible that there are unobservables that are correlated with the reform, we provide some evidence that this is not the case: parental education does not predict reform participation once the above controls are included. In the first stage of our analysis, we find that the reform is a significant predictor of educational attainment: exposure to the reform increases average educational attainment by 0.28 years for males and 0.16 years for females. These estimates are roughly similar to Meghir and Palme's (2005) original estimates.¹¹

¹¹ For the cohorts born in 1948 and in 1953, Meghir and Palme (2005) find that the compulsory school reform raised years of schooling among men by 0.252 (0.081) years and 0.339 (0.105) years for women.

Though we are not the first to study the impact of the Swedish school reform on crime, we believe that our study still makes a number of important contributions. First, we study the effect of education on crime for both males and females; most previous studies study only males.¹² Second, we have more than 30 years of crime data; for the youngest cohort in our analysis, we have crime records beginning at age 18. This allows us to look at the effect of education on crime at various stages of the life cycle. Third, our crime data is also detailed enough to allow us to look at the effect of education on various types of crimes (violent, property, and other) as well as crimes that are serious enough to warrant a sentence to prison.¹³

Our baseline estimates indicate that more schooling causes a significant decrease in criminal activity for both males and females at both the extensive and intensive margins. For males, one additional year of schooling decreases the likelihood of conviction by 2.4 percentage points (7.5 percent), the likelihood of incarceration by 1.1 percentage points (16 percent), the number of crimes by 0.40, and the number of days sentenced to prison by six. For females, one additional year of schooling significantly decreases the chance of conviction by one percentage point (11 percent) and the number of crimes by 0.09.

We also find strong evidence that these relationships persist across a number of dimensions. First, a significant negative effect of schooling on crime is seen across birth cohorts. Second, this negative effect of schooling on crime is observed across the life cycle. For instance, we see a significant negative effect on male convictions in each of the following age categories – 18-29, 30-39, and 40-49 – at both the extensive and intensive margins. Finally, for both males and females, schooling has a causal impact across crime categories. For instance, an additional

¹² Machin, Marie and Vujić (2011) present results for females in an appendix table; they do not find a significant effect of education on crime for females.

¹³ In contrast, Meghir, Palme and Schnabel (2011) only study males. They also have less crime data, such that: (i) the crime records for their youngest cohort begin at age 26, after crime typically peaks, and (ii) they do not observe crime type.

year of schooling significantly decreases the likelihood of a property crime conviction by 10 percent and a violent crime conviction by 13 percent for males.

The remainder of this paper proceeds as follows. Section 2 provides institutional details regarding the Swedish compulsory school reform. Section 3 describes the data and presents descriptive statistics. Section 4 analyzes the relationship between the reform and educational attainment (i.e. the first stage of our analysis) and evaluates the extent to which the reform is a relevant and valid instrument for education. Section 5 presents our baseline instrumental variable results of the impact of schooling on convictions and imprisonment while Section 6 considers the heterogeneity of these results across birth cohorts, over the life cycle, and across crime type. Section 7 concludes.

2. *The Swedish Compulsory School Reform*

A careful description of the Swedish compulsory school reform can be found in Marklund (1980, 1981). Detailed information can also be found in a report by the National Board of Education (1960) and in Holmlund (2007). The following brief description builds on these sources, which are recommended for further details on the topic.

In 1946, a parliamentary committee was given the task of analyzing the Swedish school system and to develop proposals and guiding principles for the future compulsory school. At this time, pupils generally went through grades 1 to 4 or 1 to 6 in a common school called *folkskolan*. In either fourth or sixth grade, more able students were selected (based on past performance) for the five-year or three- to four-year long junior-secondary school called *realskolan*. Remaining students stayed in the common school until compulsory education was completed. In most cases,

compulsory education was comprised of seven years, but in some municipalities, mainly the big cities, the minimum was eight years.

In 1948, the committee released its proposals. Their 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 advocates and the opponents of early tracking, 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. Tracking in the 9th grade, however, was later abandoned in favor of a completely comprehensive system.

The purposes underlying the proposal were among others to postpone the tracking decision to higher grades, in an effort to increase equality of opportunity, and to be able to meet the demand for junior-secondary education among the baby boom cohorts of the mid-1940s. To evaluate the appropriateness and whether the proposed nine-year comprehensive school would serve its purpose, 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 program started in 1949/1950. 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 had previously expanded their junior-secondary school to four years. 144 municipalities expressed an interest in implementing the reform. Of these, 14 municipalities

were selected for the first year of the assessment (1949/50), all of these were required to have an eight year comprehensive school already.

The National Board of Education continued with the implementation of the reform in the following years. Year by year, more municipalities joined the reform assessment program. Municipalities that wanted to take part in the reform were asked to report on their population growth, 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 in the initial year.

Apart from extending compulsory education from seven to nine years and postponing tracking, the educational reform was also pedagogical and had some affect on the curriculum. The main change to the curriculum was the introduction of English 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 and 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, and municipalities were compensated with ear-marked money from the central government for the increased costs following the expansion of mandatory education.

In 1962, the parliament decided to permanently introduce the nine-year school throughout the country. At this point, implementation became a matter for each municipality; by 1969 they were obliged to have the new comprehensive school running. As the timing was much in the hands of each municipality, the implementation was far from a randomized experiment; nevertheless, it provides a source of variation in schooling laws that allows us to investigate the causal impact of education on crime.

3. Data and Descriptive Statistics

3.1. Data Description

The data used in this paper were constructed as follows. Statistics Sweden began by drawing a 25 percent random sample from Sweden's Multigenerational Register, which includes all persons born from 1932 onwards who have lived in Sweden at any time since 1961. We restrict this sample to those born between 1943 and 1955. Very few individuals born before 1943 were involved in the school reform and we are not able to accurately classify their school reform status. After 1955, nearly all individuals were in the new compulsory school system and those who were not were living in particular, late reforming municipalities. Our 25 percent random sample is comprised of 481,656 individuals.

Longitudinal data concerning education, parish and municipality of residence in 1960 for each individual were matched on to this sample by Statistics Sweden using the unique individual identity number that each Swedish resident possesses. Parish and municipality of residence are taken from the 1960 census and are used to establish an individual's school reform status. Information on education level is taken from Sweden's Education Register and from the 1970 census. Education is recorded in 7 levels. We assign years of schooling as follows: 7 for old

primary school, 9 for new compulsory school, 11 for short high school, 12 for long high school, 14 for short university, 15.5 for long university and 19 for a Ph.D.

Sweden's official crime register was then matched on using the same personal identification number by Sweden's National Council for Crime Prevention. Thus, our data include a full record of criminal convictions for the years 1973 to 2007 for each individual in the data set; for the oldest birth cohort (born 1943), our crime data span ages 30 to 64 and for the youngest birth cohort (born 1955), our crime data span ages 18 to 52.

We use these data to construct a number of crime variables. The first variable, *crime*, is a measure of crime at the extensive margin. That is, it is equal to one if a person has ever been convicted of a crime between 1973 and 2007 and zero if she/he has not. We also identify whether a person has been convicted of a violent crime, property crime, or other type of crime between 1973 and 2007.¹⁴ At the intensive margin, we create a measure of the number of crimes that a person has been convicted of that we label *crimesum*. This variable is also broken down by crime type: violent, property and other.

One conviction may include several crimes. Our crime type variables are created by looking over all of the crimes within a single conviction.¹⁵ Speeding tickets, parking tickets and other forms of minor disturbances (ticketable offenses) are not included in our crime measure. It must be an offense that is serious enough to be taken up in court and that results in an admission of guilt or a guilty verdict.

¹⁴ Violent crimes, or crimes against persons, are crimes covered by chapters 3-7 in the Swedish criminal code (*brottsbalken*). Property crimes are those included in chapters 8-12 in the criminal code. These are standard definitions used by Sweden's National Council for Crime Prevention. All remaining crimes are labeled as "other". The 5 most common violent crimes are (in order of frequency) assault, molestation, unlawful threat, aggravated assault and aggravated unlawful threat. The five most common property crimes are petty theft (mainly shoplifting), theft, vandalism, larceny and fraud. The five most common "other" crimes are dangerous driving, driving without a license, unlawful driving, smuggling and minor narcotic offenses.

¹⁵ Thus, if you steal a car, then commit an armed robbery and then get caught after a high-speed chase, you will have one trial and one sentence that include convictions for at least three crime types. In this case, the individual would receive violent = 1 (armed robbery), property = 1 (car theft), and other = 1 (serious traffic offense + resisting arrest).

Our data also include information on the type and severity of each sentence handed down by the court. We use this information to construct a dichotomous variable, *prison*, which indicates if an individual has ever been sentenced to prison. We also create an intensive margin variable *prison sentence*, which is equal to the total number of days a person has been sentenced to prison (i.e. we sum across all known prison sentences).¹⁶

Those who died or emigrated from Sweden before 1974 are dropped from the sample, as they cannot show up in our crime data. This reduces our sample by 26,338 individuals to 455,318. We also remove those who emigrated to Sweden after 1955, since we know that these individuals did not take part in the compulsory school reform. This reduces our sample by 79,723 individuals to 375,595 individuals.¹⁷

3.2. *Determining an individual's treatment status*

The Swedish administrative registers do *not* contain information on whether individuals in the affected cohorts went through the old or new school system. However, with help from other sources, it is possible to deduct when, and for which grades, each municipality introduced the new comprehensive school. Based on this information, one can assign a reform dummy to the individuals in a data set extracted from registers.

In the registers, birth year is known, and through the censuses, it is feasible to track in which municipality an individual lived at the time of compulsory education. It is then possible to attach a reform indicator to each individual based on year of birth and municipality of residence,

¹⁶ Although we know the total sentence length, we do not know the number of days a person actually serves. In Sweden, prisoners are generally released after serving two-thirds of their sentence.

¹⁷ We have also created a larger sibling sample that includes all of the individuals in our random sample together with all of their full biological siblings born between 1943-1955. This matching is made possible by the fact that all family relationships are recorded in Sweden's Multigenerational Register. This larger sample includes 750,225 individuals, which is approximately 50 percent of all individuals in the 1943-1955 birth cohorts. None of the results reported in this paper change if we use this larger sample instead of our random sample.

maintaining the assumption that individuals are in the right grade according to their age. In some cases, it is also necessary to use 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.

There are two possible ways to construct a reform coding that can be matched to individual-level register data. One is based on existing public documents describing and evaluating the school reform. The second is deduced from register data. Following Holmlund (2007), we label them *Reform Coding 1* (based on documentation) and *Reform Coding 2* (deduced from register data).

The primary information sources necessary to construct *Reform Coding 1* are The National Board of Education (1953-1962) and Marklund (1981), which include lists of which municipalities implemented the reform each year. From these records, it is also possible to see which grades were affected in a particular municipality. These two sources cover the assessment period and only allow coding of the cohorts born 1943-1949. The remaining cohorts are coded using municipality-level tables of the number of pupils in each grade in the old and new school system published by the Educational Bureau (*Undervisningsbyrån*) (1960-1964) and Statistics Sweden (1968-1969).¹⁸

Register data sets, with large sample sizes, also allow for another procedure to code the reform, *Reform Coding 2*. For each municipality/birth year cell, it is possible to deduce the minimum level of education, and if the minimum level jumps up from *folkskola* (the old

¹⁸ Appendix A in Holmlund (2007) explains in detail how the different sources have been used to create *Reform Coding 1* 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.

compulsory minimum) to *grundskola* (the new minimum), it tells us when the reform was implemented.¹⁹

3.3. Descriptive Statistics

Descriptive statistics broken down by gender are reported in Table 1. The first three columns report statistics for the 25 percent random sample. We see that the average length of education for those born between 1943 and 1955 is 11.37 years for males and 11.72 years for females. According to *Reform Coding 1*, 39 percent of individuals were subject to the reform. According to *Reform Coding 2*, 43 percent were subject to the reform. We also see that 32 percent of all males have at least one conviction and that males have (on average) been convicted of 1.95 crimes.²⁰ Only 9 percent of the females in our random sample have been convicted of a crime. On average, females are convicted of 0.26 crimes. Lastly, we see that seven percent of males and less than one percent of females have been sentenced to prison at least once.

The last three columns in Table 1 report descriptive statistics for all individuals included in the random sample who have non-missing values for *Reform Coding 1*. The descriptive statistics of this group are nearly identical along all dimensions to those of the purely random sample.²¹

¹⁹ We use our sibling sample (see footnote 15) – which is twice as large as our random sample - when creating *Reform Coding 2*, since the algorithm works best with large samples. To create *Reform Coding 2*, we start by throwing out all individuals with education level 3 or higher. We then collapse the data into cell means by birth year and municipality. We then find the first birth year cohort within each municipality that has an average education level above 1.75. This cohort is designated as the first treated cohort. For the 3 largest cities (Malmö, Gothenburg and Stockholm) we create the reform code at the parish level instead. The correlation between *Reform Coding 1* and *Reform Coding 2* is 0.91. If we, instead, label the first treated cohort as the cohort that has an average education level above 1.50, then this correlation is 0.83. This correlation is maximized at the value of 1.75.

²⁰ While these crime rates may seem high, they are consistent with those reported in other papers using administrative Swedish crime records (see Hjalmarsson and Lindquist (forthcoming), Meghir, Palme and Schnabel (2011), and Grönqvist (2011)). It is also important to note that a majority of convictions are in the ‘other’ crimes category, which includes, for instance, a lot of alcohol related offenses.

²¹ The same is true when we look at the descriptive statistics of those who have both *Reform Coding 1* and *Reform Coding 2* defined.

Figure 1 depicts the share of individuals in each birth cohort that have been classified as treated by the compulsory school reform. Figure 2 shows the average number of years of schooling for males and females by birth cohort. Figure 3 illustrates the share of males and females in each birth cohort who have at least one conviction, as well as those with at least one prison sentence. It is important to understand that the upward trends seen in Figure 3 are driven by the fact that we have more years of crime data at younger (more crime prone) ages for the later cohorts. It is not the case that crime is simply trending upwards over time.²²

4. An Initial Analysis of the Swedish Compulsory School Reform

4.1. The Effect of the Reform on Educational Outcomes

We begin our initial analysis of Sweden's compulsory school reform by looking at the effect of the reform on educational outcomes. Did the reform raise average years of schooling and, if so, by how much? To answer these questions, we estimate the following differences-in-differences regression equation:

$$(1) \quad S_{icm} = \alpha_0 + \alpha_1 REFORM_{cm} + \alpha_2 X_{icm} + \eta_c + \mu_m + trend_m + v_{icm} ,$$

where S_{icm} is years of schooling for individual i in birth cohort c who goes to school in municipality m . $REFORM_{cm}$ is an indicator that takes the value one if the individual belongs to a birth cohort that was subject to the reform in her particular municipality. X_{icm} is a vector of observable characteristics, while η_c and μ_m represent birth cohort and municipality fixed effects, respectively. We also allow for municipality-specific time trends, $trend_m$.

²² If we create these figures using age-specific variables, e.g. whether an individual has a conviction between the ages of 30 and 39 or between the ages of 40 and 49, then all of the upward trends in crime disappear except for the trend in violent crime for men aged 30 to 39. In this figure, we see a one percentage point increase when comparing the oldest to the youngest cohorts.

The results for males and females are reported separately in Table 2. To better understand the role played by municipality fixed effects and municipality-specific trends, neither is included in the first two columns. When only birth year effects are included, we find that the average effect of the school reform on schooling is 0.64 years for males and 0.44 years for females. Controlling for parental education lowers these effects to 0.48 years and 0.31 years, respectively. Including municipality fixed effects in column (3) reduces the magnitude of these reform effects to 0.24 and 0.12 years of schooling, respectively. After adding municipality-specific linear time trends (in columns (5) and (6)), the effect of the school reform stabilizes at 0.28 years for men and 0.16 years for women. We obtain the same point estimate with and without controls for parental education. When we allow for quadratic (as opposed to linear) time trends, the estimates are slightly reduced to 0.24 for men and 0.15 for women and become less precise.²³

The gender difference in the reform effect is not surprising given the pattern of gender differences in years of schooling seen in Figure 2. Women born in these cohorts were already more likely to study beyond grade seven than men. Thus, fewer women were pushed into grade nine by the compulsory school reform.²⁴

4.2. *The Reform as an Instrument for Years of Schooling*

The purpose of this paper is to assess the causal impact of years of schooling on crime. To this end, we estimate the following baseline model:

²³ A visual inspection of the data on years of schooling at the aggregate level (see Figure 2) and at the municipality level shows that years of schooling tend to grow linearly. Any short-run accelerations or decelerations in the data are likely caused by the reform itself. Hence we prefer not to control for quadratic trends. We will, however, show how sensitive our results are to the inclusion of quadratic trends.

²⁴ The effects reported in Table 2 are most likely downward biased due to the presence of measurement error in our coding of the reform variable. However, if we use our second (independent) measure of the reform, *Reform coding 2*, as an instrument for *Reform coding 1*, then the average effect of the school reform for males rises from 0.28 (column (5)) to 0.42 and for females it goes up from 0.16 (column (5)) to 0.22. Note, however, that these are not consistent measures of the effect of the school reform. They are upwardly biased measures (see Kane et al. (1999) and Holmlund (2007) for a discussion). Thus, the true effect of the school reform lies somewhere in between.

$$(2) \quad C_{icm} = \beta_0 + \beta_1 S_{icm} + \omega_c + \theta_m + trend_m + \varepsilon_{icm}$$

where C_{icm} is a measure of criminal activity for individual i belonging to birth cohort c and going to school in municipality m . Years of schooling are given by S_{icm} and ε_{icm} is the regression error term. Similar to equation (1), equation (2) also includes birth cohort effects, ω_c , municipality fixed effects, θ_m , and municipality-specific time trends, $trend_m$.

Table 3 presents the results of estimating equation (2). We see a negative and significant relationship between years of schooling and criminal activity. This relationship is observed at both the extensive and intensive margins and for both males and females. At the extensive margin, one additional year of schooling decreases the likelihood that a male (female) has a criminal conviction by 2.3 (0.5) percentage points, on average. This relationship also exists at both margins when looking at each crime category and sentences to prison.

However, it is not necessarily the case that the coefficients presented in Table 3 can be interpreted as the causal impact of schooling on crime. The main threat to identification is that the error term includes unobservable, individual characteristics, such as cognitive and non-cognitive abilities, that may be correlated with both years of schooling and criminal behavior. This would bias our estimate of β_1 , preventing us from giving it a causal interpretation.

To address this problem, we apply an instrumental variables (IV) approach. More specifically, we will use the Swedish compulsory school reform as an instrument for years of schooling to identify the causal impact of schooling on crime, where equations (1) and (2) are the first and second stages, respectively.

A high quality instrument should be highly correlated with the variable for which it is instrumenting. Although it is not perfectly clear how one should define “highly correlated”, Staiger and Stock’s (1997) rule of thumb has gained a large measure of acceptance in the

literature. They argue that when the F -statistic on the instrument is below 10, the instrument should be considered weak. In Table 2, we see that we have only one F -statistic that is less than 10 (equal to 8). The F -statistic in our baseline model (column (5)) is 36 in the specification for females and 80 in the specification for males, implying that our instrument is, in fact, a strong predictor of years of schooling.

To be a valid instrument, it is crucial that the reform is uncorrelated with unobserved characteristics that also determine crime.²⁵ As explained in Section 2, reform implementation was not random across the population or municipalities. Instead, municipalities with certain characteristics were invited to participate in the “experiment”. Moreover, after the assessment period had come to an end, municipalities themselves decided when to join. This indeed makes it likely that the reform is correlated with characteristics specific to the municipalities. Worrisome as this may sound, it is taken care of by our differences-in-differences specification; municipality dummies control for all municipality-specific characteristics that are constant over time, while municipality time trends deal with all municipality-specific characteristics that trend over time. Nevertheless, it is still possible that the reform correlates with other unobserved factors not addressed in our analysis thus far.

A first test as to whether the reform is exogenous is already provided in Table 2. The differential results when excluding/including municipality effects clearly show that the reform is correlated with municipality-specific factors. Given the differences-in-differences specification however, if the reform is uncorrelated with unobserved characteristics, the point estimate of the effect of the reform should remain constant once we include controls for further background characteristics. Compare, for example, columns (3) and (4) in Table 2. When parental education

²⁵ Furthermore, the interpretation of our results is based on the assumption that the primary effect of the school reform on crime is through an increase in years of schooling and not through changes in the curriculum or through changes in peer groups.

is added to the specification, the point estimates are somewhat reduced, which signifies that the reform is not entirely exogenous, and that it is likely positively correlated with other factors that positively determine children's education (such as parental education). However, once we add municipality-specific linear time trends to the model, then adding parental education no longer affects the reform coefficient (compare columns (5) and (6) in Table 2).

A more direct approach to assessing whether or not the reform is correlated with parental background is to check whether or not parental education can predict reform participation after controlling for birth year effects, municipality effects and municipality-specific trends. In Table 4, we see that parental education does not predict reform participation in our baseline models shown in columns (3) and (6), which include birth year effects, municipality effects, and municipality-specific trends.

If the reform is (conditionally) exogenous, it should not have any effect on cohorts that passed through the educational system before the reform was introduced. If there is such an effect, then it is a signal that the policy is actually correlated with some unobserved factors that are not captured by our differences-in-differences approach. In Table 5, we present results for the effect of the reform on cohorts that were one, two, three, and four years too old to have been subject to any changes. What we find is that the reform did, in fact, have a sizeable, statistically significant effect on years of schooling among men and women who were one year too old to have been affected by the school reform, even after controlling for birth year and municipality effects and municipality-specific trends.

The finding of a pre-reform effect for the cohort one year ahead is not entirely unexpected. First of all, there is some measurement error in the reform coding, since in some cases it is not possible to assign a clear-cut starting date to the reform. This introduces coding

errors, but only for one cohort ahead or behind what is coded as the starting year. Second, one underlying assumption of the analysis is that individuals are in the expected grade according to their age. Some of those who repeat a grade might actually have gone to the new school, although in the data they are coded as non-participants. This should also give us a positive effect of the reform for those in the cohort one year ahead. Therefore, the fact that we only observe an effect for the cohort one year ahead, and not those two, three, and four years ahead (see Table 5), is not problematic, given that we add one pre-reform dummy to the estimated model.

As demonstrated above, measuring the reform effect is sensitive to the inclusion of municipality-specific trends. Wolfers (2006) studies the sensitivity of differences-in-differences estimates to the inclusion of region-specific trends. He argues that adding region-specific trends to the regression may capture actual responses to a policy change, and not just control for pre-policy trends. If a policy implies both a level and a trend shift, then municipality-specific trends will actually partly control for the policy responses that we want to estimate. This can lead to biased estimates, and the problem is aggravated when there are few observations before the policy change is in effect.

To examine this possibility, we re-estimated the reform effect on years of education, including municipality-specific, predicted pre-reform trends in education as a control. The trend is predicted using only observations prior to the introduction of the reform and, therefore, should not capture trend shifts induced by the policy.²⁶ The measured reform effect is, in fact, stable to

²⁶ We predict pre-reform trends in education using cohorts born 1932-1955. For individuals born 1932-1942, we use their mothers' municipality in 1960 as an indicator of where they went to school (the assumption being that mothers are less mobile than young people leaving the parental home). We first regress years of schooling on municipality and cohort dummies, and municipality-specific trends, using only the years *prior* to the implementation of the reform in each municipality. Then the predicted values from this regression are included as a control in a regression of years of schooling on the reform dummy and municipality and cohort controls, using the 1943-1955 cohorts.

the inclusion of pre-reform trends: the coefficients are close to those reported in column (3) of Table 2, around 0.13 for women and 0.24 for men.

We have also investigated post-reform dynamics. In column (1) of Table 6, we see that the effect of the school reform on years of schooling among males dies off as we examine younger cohorts born further and further away from when the reform was introduced in their municipality. Column (2) of Table 6 shows that the effect becomes insignificant (at the 5 percent level) for women born three or more years after the reform was implemented in their municipality.²⁷

A more general concern is the fact that some of our “treated” individuals were born up to 17 years after the reform was introduced in their municipality (the median is four), while some of those in our “control” group were born as many as 15 years before the reform was implemented in their municipality (the median is 5).²⁸ The fact that the compulsory school reform was implemented over such a long time period raises concerns about the comparability of the treatment and control groups. Should one really compare those born 10 years after the reform was implemented in their municipality to those who were born 10 years before the reform was implemented in their municipality? It does not seem plausible that our limited set of controls is enough to deal with the amount of unobservable heterogeneity that could exist between these two groups.

²⁷ A similar pattern is observed when using municipality specific quadratic trends rather than linear trends. Results available from the authors upon request.

²⁸ Some of the observations in our data come from the earliest reforming municipalities. The first birth cohort affected by the reform in these early reforming municipalities was the cohort born in 1938. Some of the observations in our data lived in municipalities that were late reformers. The first birth cohort affected by the reform in these late reforming municipalities was the cohort born in 1958.

Thus, to make our implicit treatment and control groups more comparable, and to deal with the reform fade-out effect seen in Table 6, we limit our sample to those individuals born at most five years before or after the reform was implemented in their municipality.²⁹

5. *The Causal Effect of an Increase in Schooling on Crime and Incarceration*

This section presents our baseline estimates of the causal effect of schooling on crime and incarceration found by estimating the two stage least squares model presented in equations (1) and (2), which uses exposure to the reform as an instrument for years of schooling. Taking the discussion from Section 4 into account, we restrict the sample to those born within five years of the first cohort affected by the reform in their own municipality. Table 7 presents the results for males (columns (1) – (3)) and females (columns (4) – (6)) for four dependant variables: crime at the extensive margin (*crime*), crime at the intensive margin (*crimesum*), prison at the extensive margin (*prison*), and prison at the intensive margin (*prison sentence*). For each dependant variable, we consider three specifications. Columns (1) and (4) include just municipality and birth cohort fixed effects while columns (2) and (5) add in municipality specific linear time trends. Columns (3) and (6), our preferred specifications, also control for a dummy indicating whether a cohort was born one year prior to the reform.

The first thing to note is that controlling for municipality specific time trends is clearly important, both qualitatively and quantitatively. In fact, without such time trends, a significant relationship is only observed between years of schooling and the number of days sentenced to

²⁹ A final threat to identification that we would like to mention is selective mobility. Geographical mobility can be a potential problem when analyzing policy changes at the regional level. Individuals may respond to the new policy, by moving away from it in order to avoid it, or by moving to it in order to benefit from it. Using comparable data sets to our own, Meghir and Palme (2003) and Holmlund (2007) have investigated this issue in some detail. Both studies conclude that there was little systematic mobility with respect to family background. Nevertheless, mobility based on unobserved individual traits might still be a problem, and we should keep this in mind when interpreting the results.

prison for males; though a number of the other point estimates are negative, they are far from significant. Controlling for linear municipality specific time trends results in a negative and significant effect of schooling on all four dependant variables for males and the extensive and intensive margin crime variables for females. While such effects are not observed for female prison sentences, it is important to recall that the proportion of females with a prison sentence is extremely small – less than one percent. Finally, the magnitude and significance of these point estimates change little when the one year pre-reform dummy is included.

We will take the results presented in columns (3) and (6) of Table 7 as our baseline estimates. We find that one additional year of schooling for males significantly decreases the likelihood of conviction by 2.4 percentage points (or by 7.5 percent) and the number of crimes by 0.40, on average. One additional year of schooling also decreases the likelihood of a prison sentence by 1.1 percentage points (or by almost 16 percent) and the number of days sentenced to prison by six days on average. For females, we observe that one additional year of schooling decreases the likelihood of conviction by one percentage point (or by 11 percent) and the number of crimes by 0.09, on average. Finally, we note that the causal instrumental variable estimates presented in Table 7 are very similar, particularly for males, to the non-causal OLS estimates in Table 3.³⁰

Overall, the results in Table 7 indicate that more schooling causes a significant decrease in criminal activity at both the extensive and intensive margins for males and females. For males, this effect is not limited to minor offenses but also has at least as large an impact on offenses serious enough to warrant a prison sentence. In addition, the magnitude of the effect of one additional year of schooling for males is comparable to that found previously in the literature; Lochner and Morretti (2004) find that one additional year of schooling decreases the likelihood

³⁰ This finding is consistent with the results of Lochner and Morretti (2004), who also find a similar pattern.

of arrest and imprisonment by about 11 percent.³¹ Given the lack of previous research, such a comparison cannot be made for females. The female effect, however, does seem to be just as large as the male effect when measured relative to the proportion of females who are convicted.

We have also estimated a number of specifications to assess the sensitivity of the results to our choice of the baseline model. First, Appendix Table 1 replicates Table 7 using *Reform Coding 2*, as described previously, to measure exposure to the reform rather than *Reform Coding 1*. The qualitative and quantitative patterns of the results do not change. Appendix Table 2 demonstrates that our results are not sensitive to the inclusion of just one pre-reform dummy; allowing individuals born two, three, and four years before the reform to be impacted by the reform does not substantively change the impact of schooling on convictions and incarceration. However, we see that significance is lost, primarily due to a large decrease in precision, when quadratic municipality specific time trends are included rather than linear trends.

6. *Heterogeneity of The Baseline Results*

6.1. *Results by Birth Cohort*

Table 8 presents the results of estimating our baseline model for three sets of birth cohorts: 1943-1947, 1947-1951, and 1951-1955.³² Are the baseline results presented above driven by any particular birth cohort? For males, we see a fairly consistent negative effect of education on

³¹ Meghir, Palme and Schnabel's (2011) estimates are consistently lower and less precise than our own. This may be partly due to the fact that we have eight more years of crime data available to us and can thus observe our individuals at younger, more crime prone ages. But this difference may also be due to the fact that we use parish and municipality at school age to assign treatment (whereas they use parish and municipality at birth), differences in our preferred specifications, the fact that we use more narrowly defined treatment and control groups, and to the fact that they report reduced form coefficients, while we report estimates from two-stage least squares. It is important to keep in mind, however, that the main focus of their paper is on the impact of parental education on children's crime. In this part of their analysis, the effects are both large and quite precisely estimated.

³² These specifications still restrict the sample to individuals born within five years of the first reform birth cohort in their own municipalities. However, the results are quite similar if we eliminate this restriction and use all individuals born in these birth cohorts. By restricting the sample to just a few birth cohorts, we are already, to some extent, controlling for differences between treatment and control groups that may exist over time.

conviction across birth cohorts. At the extensive margin, an additional year of education decreases the likelihood of a male's conviction by 2.5 percentage points if born between 1943 and 1947 or 1947 and 1951. For the youngest cohorts (1951-1955), the effect appears somewhat larger – 3.2 percentage points. A significant negative effect of education on the number of crimes is also observed for each group of birth cohorts; however, once again, there is a slightly larger effect for the youngest cohorts (-0.231) compared to the oldest cohorts (-0.166). With regards to incarceration, we see a negative coefficient for all birth cohorts at both margins; once again, this effect tends to be larger in size and more significant for the younger cohorts.

For females, we find a significant negative effect on the likelihood of conviction for cohorts born in 1943-1947 and 1951-1955. The negative effect observed at the intensive margin in Table 7 appears to be driven by the youngest birth cohorts.³³

6.2. Results Over the Life Cycle

The conviction and prison variables used in the analysis thus far consider whether an individual has a record (and the extent of the record) during the period 1973 to 2007, regardless of an individual's age. That is, for our oldest birth cohort, we are considering their criminal records from age 30 to 64; but, for our youngest birth cohort, we are considering criminal records from age 18 to 52. The use of these crime measures can raise two potential concerns. First, are any differential effects that are observed across birth cohorts (as described in the previous subsection) driven by the fact that we are measuring crime at different ages for different birth cohorts? That is, do we observe a larger effect of education on conviction and incarceration for younger male birth cohorts than older male birth cohorts in Table 8 because we can capture

³³ Note that estimates are not provided for females for the prison variables given the low intensity with which females are sentenced to prison and the reduced sample sizes in the cohort analysis, when compared to the entire sample.

criminal behavior between the ages of 18 and 30 – when criminality generally peaks – for these younger cohorts? Second, and more generally, are there differential effects of education on crime across the life cycle?

To address these questions, we create additional conviction and incarceration variables that are *age-specific*. We determine whether individuals have any convictions or prison sentences (and the number of crimes and days sentenced to prison) during four age ranges: 18-29, 30-39, 40-49, and 50-59. For all individuals in our sample, we have crime data for the middle two age categories; that is, if an individual has a record between the ages of 30 and 49, we can observe it. However, the same can only be said for the 50-59 age range for individuals born between 1943 and 1948 and for the 18-29 age range for individuals born in 1955. Table 9 presents the results of estimating our baseline specification for each age-specific crime measure at the extensive margin (columns (1) – (4)) and the intensive margin (columns (5) – (8)). With the exception of the 18-29 crime variables, these specifications are restricted to the sample of birth cohorts for whom crime records are available for all of the relevant ages. Since this is just one birth cohort for the 18-29 variables, we use the three youngest birth cohorts, who have crime records available for all ages between 20 and 29.

The results presented in the top panel of Table 9 indicate that one additional year of education significantly decreases the likelihood that a male is convicted between the ages of 18 and 29, 30 and 39, and 40 and 49; there is also a significant decrease in the number of crimes during these periods. The magnitudes of the point estimates are fairly constant across ages at both the extensive and intensive margins. But, it is important to recognize that the average conviction rate decreases with age: 28.9, 13.3, 11.3, and 8.6 percent of males have at least one conviction from 18 to 29, 30 to 39, 40 to 49, and 50 to 59, respectively. Thus, a 1.3 percentage

point effect on the likelihood of conviction between the ages of 18 and 29 (column (1)) translates into just a 4.5 percent decrease while a 1.7 percentage point effect on the likelihood of conviction between ages 40 and 49 (column (3)) translates into a 15 percent decrease. A significant effect on male convictions is not observed between the ages of 50 and 59 at the extensive or intensive margins.³⁴ These results indicate that the finding in Table 8 of a larger effect of schooling on male conviction rates for younger than older birth cohorts is not driven by the fact that we observe convictions for these cohorts between ages 18 and 29, as the estimated effect is actually smaller (in both percentage point and percent terms) for the 18-29 category than the other categories. The top panel of Table 9 also indicates a significant negative effect of an additional year of education on the likelihood of a prison sentence between ages 40 and 49 and the length of prison sentences between 30 and 39 and 40 and 49.

For females, in the lower panel of Table 9, we observe a significant negative effect of education on convictions between the ages of 30 and 39 at both the extensive and intensive margins. If we include the whole sample rather than just those with records available between the ages of 20 and 29, then we also observe significant negative effects at both margins for the 18 to 29 conviction variables. Given that less than one percent of females have a prison record at any time (and at most 0.2 percent have a record during each age period), we again do not present results for female prison sentences.

Overall, the results from Table 9 provide evidence that additional education may have a long-term impact on criminal behavior throughout the life cycle. Specifically, they indicate that the effect of an additional year of education on the likelihood of conviction and the number of convictions is fairly homogeneous across the life cycle for males but that there is a larger effect

³⁴ Note that if the entire sample is used for the age 50 to 59 regressions, then a significant effect is observed at the extensive margin for male convictions (-0.009), male prison sentences (-0.003), and female convictions (-0.003). All individuals have at least one year of conviction records available for the age 50 to 59 variables.

on prison sentences (or offenses serious enough to warrant a prison sentence) between the ages of 30 and 49. For females, the overall effects presented in Table 7 appear to largely be driven by convictions between the ages of 30 and 39.

6.3. Results by Crime Type

Table 10 considers the heterogeneity of the estimated effect across our three crime categories: property, violent, and other. As in our baseline model, all specifications include municipality and birth cohort fixed effects, municipality specific linear time trends, and one pre-reform dummy. For both males and females, a negative relationship between years of schooling and convictions is seen for all crime categories and at both the extensive and intensive margins; ten out of twelve of these estimates are significant.

More specifically, for males, we see that an additional year of schooling decreases the likelihood of a property crime conviction by 1.1 percentage points (or 10 percent), a violent crime conviction by 0.8 percentage points (or 13 percent), and a conviction of another type of crime by 1.5 percentage points (or five percent). These results are very much in line with the previous literature, as Lochner and Morretti (2004) find that an additional year of education decreases the likelihood of arrest for property and violent offenses by 11-12 percent. For females, an additional year of schooling significantly decreases the likelihood of conviction for a property offense by 1.1 percentage points (or 28 percent) and a violent offense by 0.5 percentage points (or 50 percent).

These results indicate that the negative effect of education on crime is not driven by a particular crime category and is, in fact, fairly consistent across crime categories. If anything, a smaller effect is seen for the other crimes category for both males (when compared to the

average amount of other crimes) and females (given the insignificance of the estimate). One potential explanation for this pattern is that the “other” crimes category includes numerous offenses that can be classified as white collar offenses, such as tax evasion, insider trading, forgery, anti-competitive behavior, infringement on copy- and patent-rights, neglecting work place safety rules, misconduct of public servants, etc. As suggested by Lochner (2004), a positive or, at least, less negative effect of education on crime may be expected for these offense categories.³⁵

7. Conclusion

This paper provides estimates of the causal effect of education on crime by using the change in Sweden’s compulsory schooling laws as an instrument for years of education. We find a significant negative effect of an additional year of schooling on crime that is robust in a number of dimensions. Specifically, this effect is seen for: (i) males and females, (ii) convictions and prison sentences, (iii) extensive and intensive margin measures, (iv) young and old birth cohorts, (v) crimes committed across the life cycle, and (vi) violent offenses, property offenses, and other offenses.

The magnitude of the effect is also quite substantial, and in line with previous estimates seen in the U.S. (Lochner and Morretti, 2004) and the U.K. (Machin, Marie and Vujić, 2011). Specifically, an additional year of education reduces the likelihood that a male is convicted or incarcerated by 7.5 and 16 percent, respectively. For females, one additional year of schooling decreases the likelihood of conviction by 11 percent. Importantly, these effects are just as large for violent crimes, which tend to be most costly to society, as property crimes.

³⁵ In future research, we may be able to study the effect of education on finer sub-categories of crime. But, it is uncertain whether these offenses occur often enough for effects to be precisely estimated.

These results are potentially quite important for two types of policy makers. First, when evaluating the benefits and costs of policies that increase education, it is important to consider the value of non-pecuniary benefits, such as decreased crime, and not just the effect of the additional education on wages and employment. Second, policy makers tasked with decreasing crime should consider indirect channels to doing so, rather than just focusing on the likelihood and severity of punishment.

References

- Becker, Gary (1968) "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76(2), 169-217.
- Buonanno, Paolo and Leonida Leone (2006) "Education and Crime: Evidence from Italian Regions," *Applied Economics Letters* 13(11), 709-713.
- Card, David, (1999) "The Causal Effect of Education on Earnings," in Orley Ashenfelter and David Card (eds.) *Handbook of Labor Economics* vol. 3, Elsevier, Amsterdam, 1801-1863.
- Grossman, M. and R. Kaestner (1997) "Effects of Education on Health," in J. R. Behrman and N. Stacey (eds.) *The Social Benefits of Education*, University of Michigan Press, Ann Arbor, MI, 69-123.
- Grönqvist, Hans (2011) "Youth Unemployment and Crime: New Lessons Exploring Longitudinal Register Data," SOFI Working Paper 7/2011.
- Harlow, Caroline (2003) "Education and Correctional Populations," Bureau of Justice Statistics Special Report, NCJ 195670.
- Hjalmarsson, Randi (2008) "Criminal Justice Involvement and High School Completion," *Journal of Urban Economics* 63(2), 613-630.
- Hjalmarsson, Randi and Matthew Lindquist (forthcoming) "Like Godfather, Like Son: Exploring the Intergenerational Nature of Crime," *The Journal of Human Resources*
- Holmlund, Helena (2007) "A Researcher's Guide to the Swedish Compulsory School Reform," SOFI Working Paper 9/2007.
- Jacob, Brian and Lars Lefgren (2003) "Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime," *American Economic Review* 93(5), 1560-1577.

- Kane, Thomas J., Cecilia Elena Rouse, and Douglas Staiger (1999) "Estimating Returns to Schooling When Schooling is Misreported," NBER Working Paper 7235.
- Lleras-Muney, Adriana (2005) "The relationship between Education and Adult Mortality in the United States," *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.
- Lochner, Lance (2004) "Education, Work, and Crime: A Human Capital Approach," *International Economic Review* 45(3), 811-843.
- Lochner, Lance (2008) "Education and Crime," *International Encyclopedia of Education, 3rd Edition*, Elsevier Ltd.
- Lochner, Lance (2010) "Education Policy and Crime," NBER Working paper 15894.
- Luallen, Jeremy (2006) "School's Out...Forever: A Study of Juvenile Crime, At-Risk Youths and Teacher Strikes," *Journal of Urban Economics* 59(1), 75-103.
- Machin, Stephen, Olivier Marie, and Sunčica Vujić (2011) "The Crime Reducing Effect of Education," *Economic Journal* 121, 463-484.
- Meghir, Costas and Mårten Palme (1999) "Assesing the Effect of Schooling on Earnings Using a Social Experiment in Education," Stockholm School of Economics, Scandinavian Working Papers in Economics No. 313.
- Meghir, Costas and Mårten Palme (2001) "The Effect of a Social Experiment in Education," Stockholm School of Economics, Scandinavian Working Papers in Economics No. 451.
- 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 (2011), "The Effect of Education Policy on Crime: An Intergenerational Perspective," Research Papers in Economic No. 2011:23, Department of Economics, Stockholm University.
- Oreopoulos, Philip and Kjell G. Salvanes (2009) "How Large are Returns to Schooling? Hint: Money Isn't Everything," NBER Working Paper No. 15339.
- Petit, Becky and Bruce Western (2004) "Mass Imprisonment and the Life Course: Race and Class Inequality in U.S. Incarceration," *American Sociological Review*, 69(2), 151-169.

Figure 1. Share of Each Birth Cohort in the Random Sample who Are Treated by the Compulsory School Reform, *Reform Coding 1* and *Reform Coding 2*.

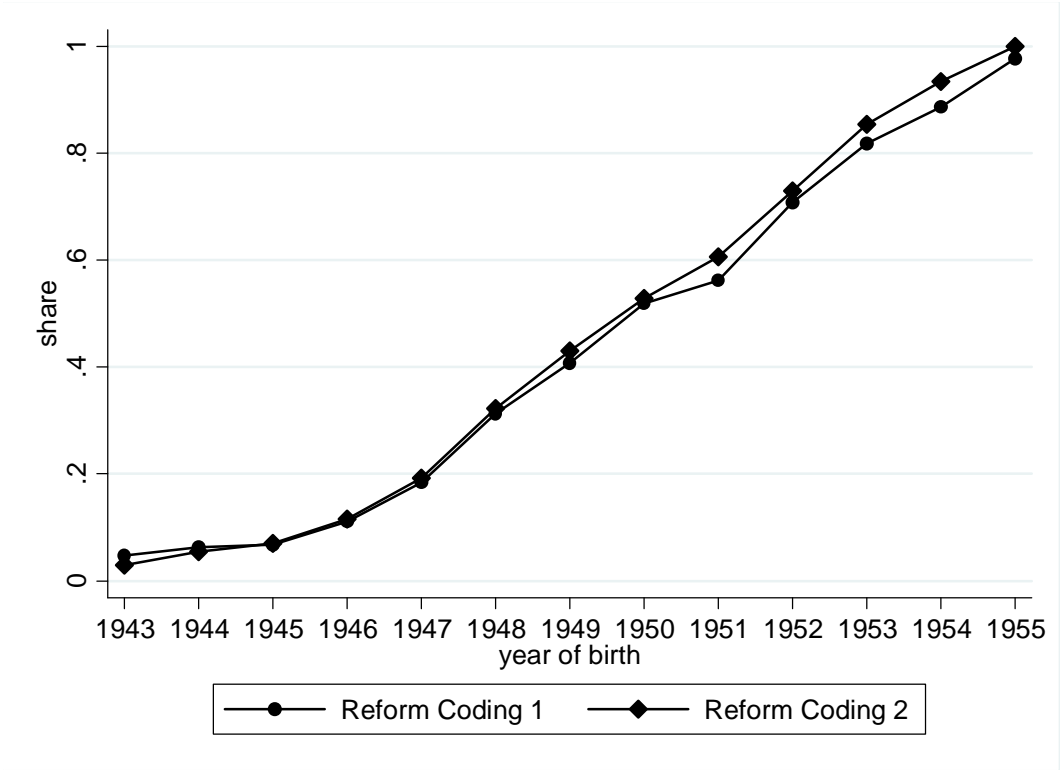


Figure 2. Average Years of Schooling by Year of Birth for Males and Females with Reform Status Defined.

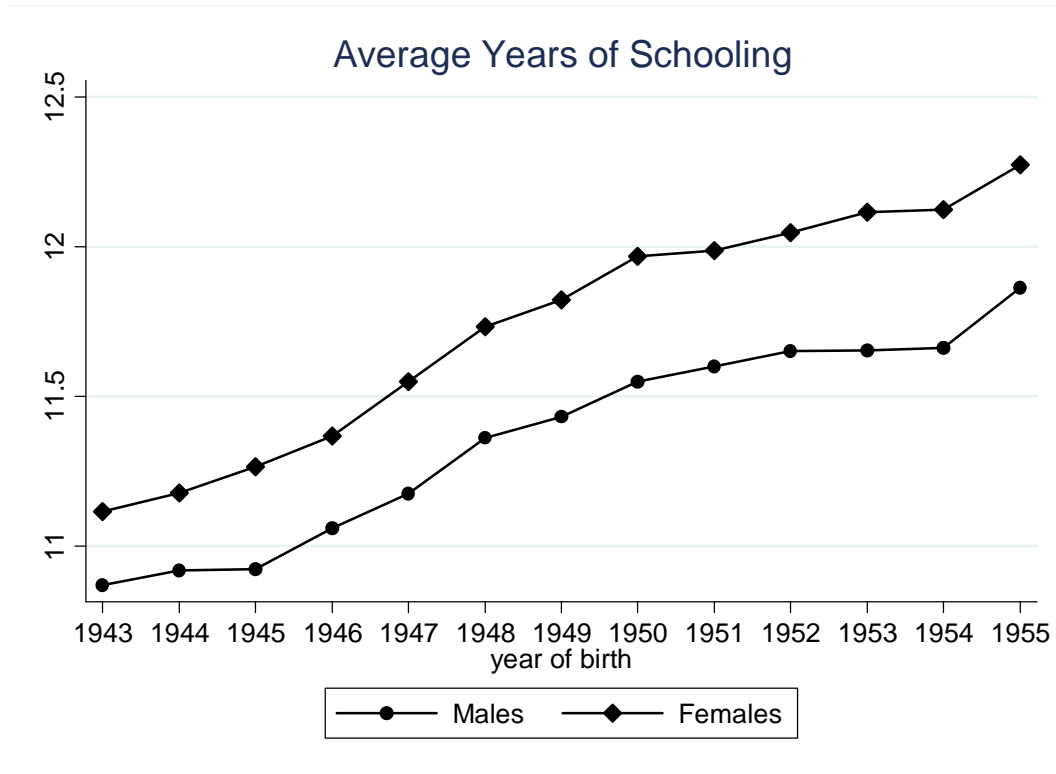


Figure 3. Share of Males and Females with Reform Status Defined Who Have at least One Conviction or at least One Prison Sentence.

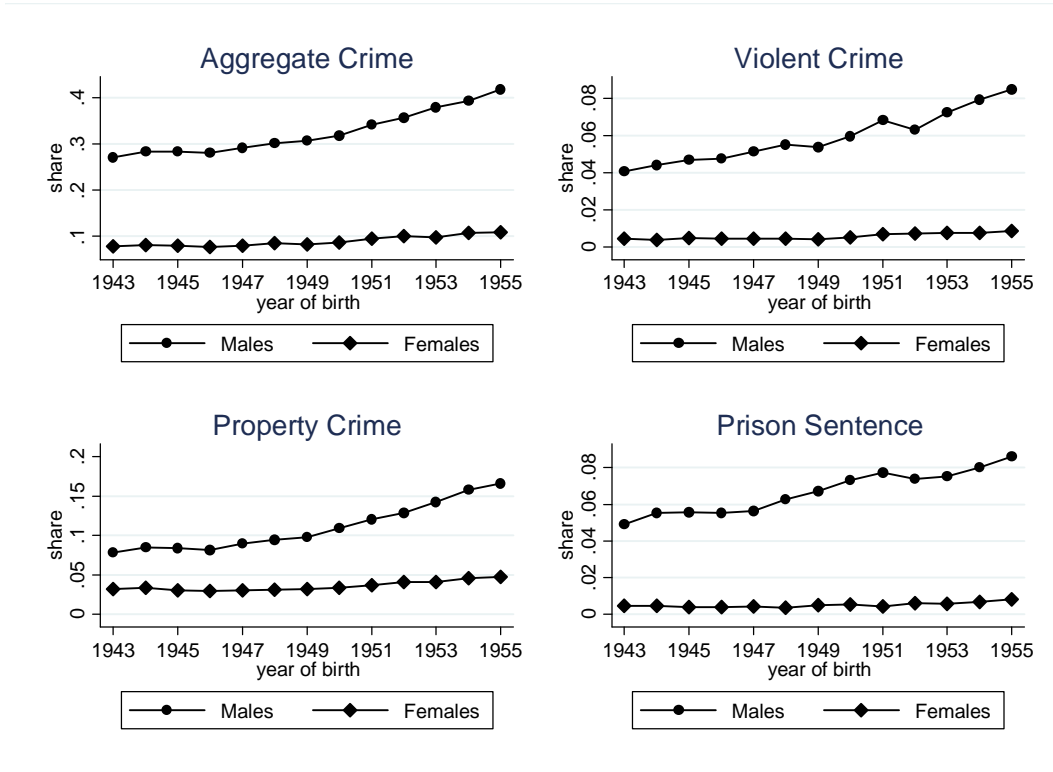


Table 1. Descriptive Statistics.

	Random sample			Random sample with <i>Reform coding 1</i> defined		
	Mean	(S.D.)	Obs.	Mean	(S.D.)	Obs.
<i>Males</i>						
<i>Birth year</i>	1949	(3.72)	192468	1949	(3.73)	168982
<i>Years of schooling</i>	11.37	(2.79)	190746	11.31	(2.81)	167813
<i>Income</i>	12.11	(0.59)	192226	12.12	(0.56)	168889
<i>Reform coding 1</i>	0.39	(0.49)	168982	0.39	(0.49)	168982
<i>Reform coding 2</i>	0.43	(0.49)	160927	0.43	(0.49)	160653
<i>Crime (yes=1, no=0)</i>	0.32	(0.47)	192468	0.32	(0.47)	168982
<i>Property crime</i>	0.11	(0.31)	192468	0.11	(0.31)	168982
<i>Violent crime</i>	0.06	(0.24)	192468	0.06	(0.23)	168982
<i>Other crime</i>	0.28	(0.45)	192468	0.28	(0.45)	168982
<i>Crime sum</i>	1.95	(10.83)	192468	1.85	(10.41)	168982
<i>Property crime sum</i>	0.63	(5.03)	192468	0.60	(4.85)	168982
<i>Violent crime sum</i>	0.17	(1.24)	192468	0.17	(1.22)	168982
<i>Other crime sum</i>	1.14	(6.14)	192468	1.09	(5.90)	168982
<i>Prison (yes=1, no=0)</i>	0.07	(0.25)	192468	0.07	(0.25)	168982
<i>Prison sentence (days)</i>	25.93	(247.74)	192468	23.99	(235.81)	168982
<i>Females</i>						
<i>Birth year</i>	1949	(3.71)	183127	1949	(3.72)	160822
<i>Years of schooling</i>	11.72	(2.60)	181882	11.69	(2.61)	159958
<i>Income</i>	11.72	(0.61)	182744	11.72	(0.59)	160601
<i>Reform coding 1</i>	0.39	(0.49)	160822	0.39	(0.49)	160822
<i>Reform coding 2</i>	0.43	(0.49)	153177	0.49	(0.49)	152938
<i>Crime (yes=1, no=0)</i>	0.09	(0.29)	183127	0.09	(0.28)	160822
<i>Property crime</i>	0.04	(0.19)	183127	0.03	(0.18)	160822
<i>Violent crime</i>	0.01	(0.08)	183127	0.01	(0.07)	160822
<i>Other crime</i>	0.06	(0.24)	183127	0.06	(0.24)	160822
<i>Crime sum</i>	0.26	(2.69)	183127	0.24	(2.35)	160822
<i>Property crime sum</i>	0.11	(1.39)	183127	0.10	(1.25)	160822
<i>Violent crime sum</i>	0.01	(0.21)	183127	0.01	(0.21)	160822
<i>Other crime sum</i>	0.14	(1.63)	183127	0.13	(1.38)	160822
<i>Prison (yes=1, no=0)</i>	0.01	(0.07)	183127	0.00	(0.07)	160822
<i>Prison Sentence (days)</i>	1.56	(50.49)	183127	1.40	(48.16)	160822

Table 2. The Effect of the Reform on Years of Schooling.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Males</i>								
Reform effect	0.635*** [0.051]	0.477*** [0.033]	0.237*** [0.043]	0.215*** [0.052]	0.282*** [0.032]	0.281*** [0.036]	0.232*** [0.037]	0.238*** [0.039]
F-stat reform	158	206	31	17	80	62	38	37
Observations	167813	167813	167813	167813	167813	167813	167813	167813
R-squared	0.02	0.12	0.08	0.15	0.08	0.16	0.09	0.17
<i>Females</i>								
Reform effect	0.440*** [0.037]	0.310*** [0.025]	0.124*** [0.031]	0.104*** [0.036]	0.162*** [0.027]	0.162*** [0.025]	0.151*** [0.034]	0.148*** [0.030]
F-stat reform	138	153	16	8	36	42	20	24
Observations	159958	159958	159958	159958	159958	159958	159958	159958
R-squared	0.03	0.11	0.06	0.12	0.06	0.13	0.07	0.14
Year effects	x	x	x	x	x	x	x	x
Municip. effects			x	x	x	x	x	x
Parental educ.		x		x		x		x
Linear trends					x	x		
Quadratic trends							x	x

OLS regression coefficients. The dependent variable in each specification is years of schooling. Robust standard errors clustered at the municipality level are reported in brackets; *** indicates significance at 1% level.

Table 3. OLS (non-causal) Relationship between Years of Schooling and Extensive/Intensive Margin Conviction and Prison Measures

	(1)	(2)	(3)	(4)	(5)
Extensive margin variable	Crime	property crime	violent crime	other crime	prison
Years of schooling <i>Males (N = 167,813)</i>	-0.023*** [0.001]	-0.014*** [0.001]	-0.009*** [0.000]	-0.020*** [0.001]	-0.009*** [0.000]
years_of_schooling <i>Females (N = 159,958)</i>	-0.005*** [0.000]	-0.004*** [0.000]	-0.001*** [0.000]	-0.003*** [0.000]	-0.001*** [0.000]
Intensive margin Variable	crime sum	property crime sum	violent crime sum	other crime sum	prison sentence (days)
years_of_schooling <i>Males (N = 167,813)</i>	-0.334*** [0.025]	-0.126*** [0.009]	-0.031*** [0.002]	-0.177*** [0.015]	-4.796*** [0.424]
years_of_schooling <i>Females (N = 159,958)</i>	-0.039*** [0.003]	-0.021*** [0.002]	-0.002*** [0.000]	-0.016*** [0.002]	-0.347*** [0.048]

OLS regression coefficients. Robust standard errors in brackets, clustered by municipality; *** indicates significance at 1% level. All specifications include birth cohort and municipality fixed effects, as well as municipality specific time trends.

Table 4. Can Parental Education Predict School Reform Participation?

	(1)	(2)	(3)	(4)	(5)	(6)
	Males			Females		
Father education	0.008*** [0.002]	0.001* [0.001]	-0.000 [0.000]	0.009*** [0.002]	0.001** [0.000]	-0.000 [0.000]
Father ed. missing	-0.007* [0.004]	-0.001 [0.002]	-0.001 [0.001]	-0.006 [0.004]	0.001 [0.002]	-0.000 [0.002]
Mother education	0.006*** [0.002]	0.000 [0.001]	0.000 [0.000]	0.006*** [0.001]	0.000 [0.000]	0.000 [0.000]
Mother ed. missing	-0.005 [0.003]	0.000 [0.002]	-0.001 [0.001]	-0.009** [0.004]	-0.000 [0.002]	0.000 [0.002]
Year effects	x	x	x	x	x	x
Municipality effects		x	x		x	x
Linear trends			x			x
Observations	167813	167813	167813	159958	159958	159958
R-squared	0.40	0.67	0.76	0.39	0.67	0.75

OLS regression coefficients. Dependent variable is *Reform Coding 1*. Robust standard errors clustered at the municipality level are reported in brackets; *** indicates significance at 1% level; ** indicates significance at 5% level.

Table 5. The Effect of the Reform on Cohorts Who Were Too Old to Have Been Affected.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Males				Females			
Reform effect	0.336***	0.274***	0.283***	0.276***	0.188***	0.162***	0.164***	0.163***
	[0.034]	[0.033]	[0.032]	[0.032]	[0.031]	[0.030]	[0.026]	[0.027]
F-stat on reform effect	100	69	76	73	37	29	38	37
Reform effect for cohort t years too old								
$t=1$	0.118***				0.056*			
	[0.031]				[0.029]			
$t=2$		-0.022				-0.002		
		[0.028]				[0.026]		
$t=3$			0.003				0.006	
			[0.029]				[0.028]	
$t=4$				-0.041				0.004
				[0.028]				[0.026]

OLS regression coefficients. Dependent variable is *years of schooling*. Controls include year of birth and municipality effects. Municipality-specific linear trends are also included. Standard errors clustered at municipality level are reported in brackets; *** indicates significance at 1% level; ** indicates significance at 5% level; * indicates significance at 10% level.

Table 6. Post-reform Dynamics.

	(1)	(2)
	Linear trends	
	Males	Females
Reform year	0.298*** [0.045]	0.142*** [0.040]
Post-reform 1	0.333*** [0.054]	0.094* [0.051]
Post-reform 2	0.266*** [0.063]	0.130** [0.060]
Post-reform 3	0.267*** [0.077]	-0.050 [0.073]
Post-reform 4	0.229** [0.094]	-0.080 [0.096]
Post-reform 5	0.213* [0.109]	0.016 [0.105]
Post-reform 6	0.225* [0.132]	-0.123 [0.125]
Post-reform 7	0.163 [0.152]	-0.103 [0.146]
Post-reform 8	0.192 [0.175]	-0.250 [0.176]
Post-reform 9	0.130 [0.189]	-0.293 [0.200]
Post-reform 10+	0.128 [0.235]	-0.236 [0.242]

OLS regression coefficients. Dependent variable is *years of schooling*. Controls include year of birth and municipality effects. Municipality-specific linear trends are also included. Standard errors clustered at municipality level are reported in brackets; *** indicates significance at 1% level; ** indicates significance at 5% level; * indicates significance at 10% level.

Table 7. Baseline Instrumental Variable Results

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dependent Variable = crime</i>						
years_of_schooling	0.008 [0.021]	-0.023*** [0.007]	-0.024*** [0.007]	-0.01 [0.020]	-0.009** [0.004]	-0.010** [0.004]
<i>Dependent Variable = crimesum</i>						
years_of_schooling	-0.527 [0.417]	-0.324** [0.150]	-0.404*** [0.147]	-0.224 [0.179]	-0.084*** [0.021]	-0.093*** [0.021]
<i>Dependent Variable = prison</i>						
years_of_schooling	-0.013 [0.012]	-0.011*** [0.004]	-0.011*** [0.004]	0.002 [0.005]	-0.0003 [0.001]	-0.0003 [0.001]
<i>Dependent Variable = prison sentence (days)</i>						
years_of_schooling	-16.653* [9.101]	-4.707* [2.740]	-5.992** [2.752]	-3.239 [4.345]	0.322 [0.396]	0.341 [0.393]
Observations	110650	110650	110650	105877	105877	105877
Sex	Male	Male	Male	Females	Females	Females
Municipality and BC FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Linear Trends	No	Yes	Yes	No	Yes	Yes
1 PreReform Dummy	No	No	Yes	No	No	Yes

Robust standard errors clustered at the municipality level are in brackets. *** indicates significance at 1% level; ** indicates significance at 5% level; * indicates significance at 10% level. All specifications are restricted to the sample of individuals born within five years of the reform in their own municipality.

Table 8. Heterogeneity of IV Results Across Birth Cohorts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Crime			crimesum			prison			Prison sentence (days)		
<i>Males</i>												
years_of_schooling	-0.025***	-0.025***	-0.031***	-0.166**	-0.209**	-0.231**	-0.003	-0.007*	-0.015***	-3.529	-3.332	-7.439***
	[0.009]	[0.007]	[0.006]	[0.083]	[0.097]	[0.110]	[0.004]	[0.004]	[0.004]	[2.776]	[2.452]	[2.653]
Observations	32696	51119	46562	32696	51119	46562	32696	51119	46562	32696	51119	46562
<i>Females</i>												
years_of_schooling	-0.015***	-0.004	-0.010**	0.157	0.008	-0.081***						
	[0.005]	[0.004]	[0.005]	[0.128]	[0.045]	[0.020]						
Observations	31398	49074	44312	31398	49074	44312						
Birth Cohorts	1943-47	1947-51	1951-55	1943-47	1947-51	1951-55	1943-47	1947-51	1951-55	1943-47	1947-51	1951-55
Year and Municipality FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Municipality Specific	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Trend	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Pre-reform dummy	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Robust standard errors clustered at the municipality level are in brackets. *** indicates significance at 1% level; ** indicates significance at 5% level; * indicates significance at 10% level. All specifications are restricted to the sample of individuals born within five years of the reform in their own municipality. Note that, given the extremely small percentage of females sentenced to prison within a single birth cohort, the results for females are not presented when using *prison* or *prison sentence* as the dependent variable.

Table 9. Heterogeneity of IV Results Across the Life Cycle

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Extensive Margin				Intensive Margin			
	18-29	30-39	40-49	50-59	18-29	30-39	40-49	50-59
<i>Males</i>								
<i>Crime and Crimesum</i>								
years_of_schooling	-0.013** [0.007]	-0.016*** [0.005]	-0.017*** [0.005]	-0.003 [0.005]	-0.138** [0.062]	-0.109** [0.050]	-0.120*** [0.037]	-0.011 [0.019]
<i>Prison and Prison sentence (days)</i>								
years_of_schooling	-0.003 [0.003]	0.002 [0.002]	-0.005*** [0.002]	-0.001 [0.001]	-0.361 [0.965]	-3.013*** [1.027]	-1.574** [0.802]	1.389 [1.132]
Observations	25012	110650	110650	42662	25012	110650	110650	42662
<i>Females</i>								
<i>Crime and Crimesum</i>								
years_of_schooling	-0.003 [0.004]	-0.007*** [0.002]	-0.001 [0.003]	-0.001 [0.002]	-0.032 [0.029]	-0.035*** [0.008]	-0.014 [0.010]	-0.00007 [0.003]
Observations	23773	105877	105877	41028	23773	105877	105877	41028

Robust standard errors clustered at the municipality level are in brackets. *** indicates significance at 1% level; ** indicates significance at 5% level; * indicates significance at 10% level. All specifications include birth cohort and municipality fixed effects, municipality specific linear trends and one pre-reform dummy. All specifications are restricted to the sample of individuals born within five years of the reform in their own municipality. In addition, the samples are restricted to those individuals for whom with 10 years (or more) of crime data are available for each period. So, the entire sample is used for crime variables measured between 30 and 39 and 40 and 49. But just those born between 1953 and 1955 are used for the age 20 to 29 crime variables and those born between 1943 and 1948 are used for the age 50 to 59 variables. The dependent variable for each specification can be determined by looking at the panel title and the age label in each column. Thus, the dependent variable in the first cell is whether there is any conviction (*crime*) between the ages of 18 and 29.

Table 10. Heterogeneity of IV Results Across Crime Types

	(1) property crime	(2) violent crime	(3) other crime	(4) property crimesum	(5) violent crimesum	(6) other crimesum
<i>Males</i>						
years_of_schooling	-0.011** [0.004]	-0.008** [0.004]	-0.015** [0.007]	-0.091 [0.066]	-0.036** [0.015]	-0.278*** [0.088]
Observations	110650	110650	110650	110650	110650	110650
<i>Females</i>						
years_of_schooling	-0.011*** [0.003]	-0.005*** [0.001]	-0.004 [0.004]	-0.045*** [0.014]	-0.007*** [0.002]	-0.041*** [0.012]
Observations	105877	105877	105877	105877	105877	105877

Robust standard errors clustered at the municipality level are in brackets. *** indicates significance at 1% level; ** indicates significance at 5% level; * indicates significance at 10% level. All specifications include municipality and birth cohort fixed effects, municipality specific linear time trends, and one pre reform dummy and restrict the sample to those born within 5 years of the reform in their municipality.

Appendix Table 1. Baseline IV Results for Males and Females Using Reform Coding 2

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dependent Variable = crime</i>						
years_of_schooling	-0.019 [0.016]	-0.033*** [0.008]	-0.033*** [0.007]	0.007 [0.017]	-0.009** [0.004]	-0.010** [0.004]
<i>Dependent Variable = crimesum</i>						
years_of_schooling	-0.433 [0.349]	-0.355* [0.184]	-0.369** [0.181]	-0.172 [0.153]	-0.115*** [0.035]	-0.120*** [0.034]
<i>Dependent Variable = prison</i>						
years_of_schooling	-0.021** [0.009]	-0.011*** [0.004]	-0.011*** [0.004]	0.002 [0.004]	-0.001 [0.001]	-0.001 [0.001]
<i>Dependent Variable = prison sentence (days)</i>						
years_of_schooling	-13.450* [7.847]	-3.813 [3.396]	-3.979 [3.363]	-5.332 [3.732]	-0.053 [0.465]	-0.01 [0.442]
Observations	107409	107409	107409	102823	102823	102823
Sex	Male	Male	Male	Females	Females	Females
Municipality and BC FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Linear Trends	No	Yes	Yes	No	Yes	Yes
1 PreReform Dummy	No	No	Yes	No	No	Yes

Robust standard errors clustered at the municipality level are in brackets. *** indicates significance at 1% level; ** indicates significance at 5% level; * indicates significance at 10% level. All specifications are restricted to the sample of individuals born within five years of the reform in their own municipality.

Appendix Table 2. Robustness to Pre-Reform Dummies and Quadratic Time Trend

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Males						Females					
<i>Dependant Variable = Crime</i>												
years_of_schooling	-0.024*** [0.007]	-0.023*** [0.007]	-0.026*** [0.007]	-0.026*** [0.007]	-0.028*** [0.006]	-0.02 [0.036]	-0.010** [0.004]	-0.009** [0.004]	-0.009** [0.004]	-0.009** [0.004]	-0.008** [0.004]	-0.004 [0.025]
<i>Dependant Variable = Crimesum</i>												
years_of_schooling	-0.404*** [0.147]	-0.324** [0.150]	-0.425*** [0.148]	-0.426*** [0.147]	-0.438*** [0.152]	-0.692 [0.711]	-0.093*** [0.021]	-0.084*** [0.021]	-0.085*** [0.021]	-0.085*** [0.021]	-0.080*** [0.019]	-0.333 [0.245]
<i>Dependant Variable = prison</i>												
years_of_schooling	-0.011*** [0.004]	-0.011*** [0.004]	-0.012*** [0.003]	-0.012*** [0.003]	-0.012*** [0.003]	-0.015 [0.018]	-0.0003 [0.001]	-0.0003 [0.001]	-0.001 [0.001]	-0.001 [0.001]	-0.0005 [0.001]	0.008 [0.006]
<i>Dependant Variable = prison sentence (days)</i>												
years_of_schooling	-5.992** [2.752]	-4.707* [2.740]	-5.456** [2.615]	-5.226** [2.598]	-5.234* [2.730]	-23.134 [16.543]	0.341 [0.393]	0.322 [0.396]	0.481 [0.356]	0.497 [0.352]	0.605* [0.350]	-0.237 [5.507]
Observations	110650	110650	110650	110650	110650	110650	105877	105877	105877	105877	105877	105877
Municipality and Birth Cohort Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Municipality Specific Linear Time Trend	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	NO
Municipality Specific Quadratic Time Trend	NO	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO	YES
# of Pre-Reform Dummies	1	0	2	3	4	1	1	0	2	3	4	1

Robust standard errors clustered at the municipality level are in brackets. *** indicates significance at 1% level; ** indicates significance at 5% level; * indicates significance at 10% level. All specifications are restricted to the sample of individuals born within five years of the reform in their own municipality.