Institute for Fiscal Studies

The Effect of Education Policy on Crime: An International Perspective

IFS Working Paper W11/11

Costas Meghir
Mårten Palme
Marieke Schnabel

The Effect of Education Policy on Crime: An Intergenerational Perspective*

Costas Meghir†, Mårten Palme‡, Marieke Schnabel§

May 2011

Abstract

The Swedish comprehensive school reform implied an extension of the number of years of compulsory school from 7 or 8 to 9 for the entire nation and was implemented as a social experiment by municipality between 1949 and 1962. A previous study (Meghir and Palme, 2005) has shown that this reform significantly increased the number of years of schooling as well as labor earnings of the children who went through the post reform school system, in particular for individuals originating from homes with low educated fathers. This study estimates the impact of the reform on criminal behavior: both within the generation directly affected by the reform as well as their children. We use census data on all born in Sweden between 1945 and 1955 and all their children merged with individual register data on all convictions between 1981 and 2008. We find a significant inverse effect of the reform on criminal behavior of men and on sons to fathers who went through the new school system.

*We thank Matthew Lindquist, Imran Rasul, Lisa Jönsson, Emilia Simeonova as well as participants at seminars at Stockholm University, University College London, Yale University and the Institute for Labor Market Policy Evaluation (IFAU) in Uppsala for helpful comments and suggestions. We gratefully acknowledge financial support from IFAU. Costas Meghir thanks the ESRC for funding under the Professorial Fellowship RES-051-27-0204 and under the ESRC Centre at the IFS ESRC RES-544-28-5001.

†Department of Economics, Yale University; Institute for Fiscal Studies and Department of Economics, University College London, Gower Street, London WC1E 6BT, UK. E-mail: C.Meghir@ucl.ac.uk

‡Department of Economics, Stockholm University, SE - 10691 Stockholm, Sweden. E-mail: Marten.Palme@ne.su.se

§Department of Economics, University College London, Gower Street, London WC1E 6BT, UK. E-mail: Marieke.Schnabel@ucl.ac.uk

1 Introduction

Crime imposes huge costs on society and policies to prevent crime are debated around the world. In addition to police resources and punishment of criminals, these policies typically involve public spending on creating alternative labor outcomes for individuals potentially involved in criminal activities, e.g. reforms of national education systems. Since criminal behavior has an exceptionally strong intergenerational link, such policies may have an effect across generations. Although there is a very large previous literature showing that low educated people are more likely to be convicted and that their children are more likely to commit crimes, very few studies have attempted to show a causal relation between own education and criminal behavior and (to our knowledge) no previous study has shown a causal relation between parental education and criminal behavior.

In this paper we study to what extent a major national education reform, which increased schooling and earnings, also affected criminal behavior: Both in the generation directly affected by the reform as well as their children. The reform in question is the Swedish comprehensive school reform, which included an extension of the number of years of compulsory schooling and was implemented gradually across the 1,055 Swedish municipalities during the 1950s, before finally being rolled out nationally in 1962. As described in Meghir and Palme (2005), the implementation of the reform allows us to investigate its impact on crime by comparing individuals living and working in the same labour market but having been exposed to different schooling systems. Moreover, we can also estimate the impact of the reform on their own children. In this case we are able to consider how changing the educational experience of the parents can affect important outcomes for their children. Beyond the obvious public policy implications, the results speak to the growing literature

on the importance of parenting quality on child outcomes.

Two earlier papers by Lochner and Moretti (2004) and Machin, Marie, and Vujić (2010) respectively study the relation between compulsory schooling laws and criminal behavior. Lochner and Moretti (2004) use changes in compulsory schooling laws across time between US states to identify the effect of increasing education on crime. Machin, Marie, and Vujić (2010) compare criminal behavior of the cohorts just before and just after the implementation of the comprehensive school system in Britain. Our paper goes further in two important ways. First, we compare the criminal behavior of two groups, distinguished by the school system they were exposed to, but active in the same labor markets at the same point in time, and who belong to the same cohort. As a result any general equilibrium effects of the reforms are eliminated, as they are common across groups - what we identify is the impact of being exposed to alternative educational systems, over and above the macroeconomic effects that result from this. However, the key distinguishing feature of our paper is that we consider the effects on the criminal activity of the children of those affected by the reform. As we discuss later, the effects in this case will operate through improved parenting and investments in children.

The paper is organized as follows: Section 2 discusses previous theoretical and empirical work on the relation between both own education and criminal behavior as well as parental education and criminal behavior; Section 3 provides an overview of the comprehensive school reform in Sweden; Section 4 describes the data and presents empirical results on the relation between educational attainments and criminal behavior as well as intergenerational associations of crime; Section 5 discusses our identification strategy; Section 6 presents our empirical results on the effect of the comprehensive school reform and the probability of being convicted for crime or

sentenced to prison; Section 7 concludes.

2 The Impact of Education on Crime

2.1 The Impact of Education on Crime within a Generation

The links between economic incentives and crime have been established both theoretically and empirically in earlier studies. A prominent example is Freeman (1999) who outlines an economic model of crime where the choice between criminal and legal activity is determined by comparing the expected utility of each. Grogger (1998), Gould, Weinberg, and Mustard (2002) and Machin and Meghir (2004) demonstrate the importance of wages and labour market opportunities in driving crime. One implication of this is that improved education may reduce crime. A number of papers have looked at the link between education and crime directly.

These include Lochner and Moretti (2004) and Machin, Marie, and Vujić (2010), cited above. A more theoretically based approach was offered by Lochner (2004) who develops a life cycle model of education and crime and estimates a negative education-crime relationship. A study, based on this capital approach by Williams and Sickles (2002) finds that years of schooling has a significant negative effect on crime in adulthood, and that there is a relationship between crime and other measures of human capital. Earlier studies support this empirical evidence on the education-crime relationship; for example Freeman (1996) states for the 1991 US Census that two thirds of US prison inmates are high-school drop-outs and 12 percent of 24-35 year old high school drop outs were incarcerated in 1993. This negative correlation between crime and education has also been documented in the the criminology and sociology literature, for example Sabates and Feinstein (2008a).[1]

[1]See also Sabates and Feinstein (2008b).

Gallipoli and Fella (2006) estimate a general equilibrium model of crime and education. They find that increases in education have a significant impact in reducing crime. However, they point out that the general equilibrium effects, operating through changes in wages as the number of educated individuals increases, can be substantial.

2.2 The Effects of Parental Education on Children's Criminal Behavior

Intergenerational associations of criminal behavior have been documented in the criminology literature. From the economics point of view this question relates to the investments that parents make on their children and the way that parental education may affect such investments (see Becker (1981)).[2]

Of particular interest to our study are the two recent papers by Hjalmarsson and Lindquist (2010a) and Hjalmarsson and Lindquist (2010b). The first documents a strong correlation between crime of fathers and children of both genders in Sweden using the Stockholm Birth Cohort Study, which is a data set that consists of all individuals born in 1953 and living in Stockholm in November 1963, containing crime records of the birth cohort and their fathers. The second study focuses on parent-child correlations in crime using adoption data. In particular they aim to determine through which factors mothers and fathers influence child criminality, which is closely related to the study by Björklund, Lindahl, and Plug (2006).

[2]For some empirical work see for example Carneiro, Meghir, and Parey (2007). Moreover, there is direct evidence that better childhood environments and early education can reduce crime rates, see for example the results form the Perry pre-school experiment presented in Schweinhart, Montie, Xiang, Barnett, Belfield, and Nores (2005) and Cunha and Heckman (2007).

2.3 Human Capital and participation in Crime - a simple framework

To better understand the mechanisms through which educational reform can affect participation in crime for both generations consider the following simple model. Human capital is produced by investments in various stages of the child's life as well as by overall educational attainment E_c. Suppose there are two stages, early investments I_0 and investments during schooling I_1. The efficiency of investments depends on the educational level of the parent, E_p. Denote the human capital production function by

$$H = H(I_0, I_1, E_c|E_p)$$

where $H'_{I_0} > 0$, $H'_{I_1} > 0$, $H'_{E_c} > 0$ and $H''_{I_0} < 0$, $H''_{I_1} < 0$, $H''_{E_c} < 0$. Parents are assumed to care about child quality, which here is just their human capital. Ignoring dynamics for simplicity, they solve the problem[3]

$$\max_{C,I_0,I_1} \{u(C,H) \text{ st } C + I_0 + I_1 = Y^p \text{ and } H = H(I_0, I_1, E_c|E_p)\}$$

where C is parental consumption. In this simple context investments in children will increase as parental resources Y^P increase, so long as H is a normal good. The first order conditions for investments are

$$u'_H \frac{\partial H}{\partial I_0} = u'_C$$

$$u'_H \frac{\partial H}{\partial I_1} = u'_C$$

An increase in the marginal productivity of such investments (say due to an increase

[3]The problem is dynamic sequential, but nothing would be gained in introducing this notation here.

in parental education E_p) will lead to more investments in the children at both stages. This will happen both because the productivity of investments may increase and because parental resources Y^P go up. The next step is to see how these changes can affect participation in crime.

Consider a very simple model of crime choice in two stages of life. First, is the educational stage, where the individual can either engage in education, crime or work. Then follows the *adult* stage where there is no education choice. We start by the latter.

Committing a crime is a period by period decision with no dynamics (for simplicity). Working leads to income $Y(H)$ which depends on human capital H. Crime on the other hand yields a return R with some probability $p(H)$. Being caught, with probability $1-p(H)$ leads to punishment $K(H)$. As discussed in Lochner and Moretti (2004), the dependence of K on H represents the opportunity cost of being incarcerated. Moreover, it also costs $c(H)$ to participate. This cost can reflect the aversion that one may have to anti-social behaviour. We assume that $c(H)$ is increasing in H.

Participation in crime is determined by the condition

$$p(H)R-(1-p(H))K(H)-c(H)-Y(H)>0 \Longleftrightarrow \text{ engage in crime}$$

An increase in H will increase earnings $Y(H)$ and participation costs $c(H)$, both implying a reduction in crime. A possible mitigating effect is that better human capital may make crime more effective and reduce the probability of capture $p(H)$. In our empirical analysis we only measure convictions; we assume that a reduction in convictions reflects a reduction in crime participation and not more effective

criminals. Thus, other than the potential effect on $p(H)$, increasing human capital will decrease participation in crime.

It is also useful to consider the earlier period, when the individual still has the option of being in school. Define the future value as $V(H_1) = E \max_{\kappa,w}(V^{\kappa}(H_1), V^{w}(H_1))$ with V^{κ} denoting the value of crime and V^{w} the value for work. $c^{ed}(H_0)$ denotes the cost of education, which we assume are declining in H_0 (initial human capital). In this first period the value of education, crime and work respectively are given by

$$V^{ed} = -c^{ed}(H_0) + \beta V(H_1^+)$$

$$V^{w} = Y(H_0) + \beta V(H_0)$$

$$V^{k} = p(H_0)R - (1 - p(H_0))K(H_0) - c(H_0) + \beta V(H_0)$$

where H_1^+ denotes that education allows the individual to enter the next period with higher human capital. The individual choose the activity with the greatest value. First, note that if schooling is compulsory, then there is a mechanical reduction in crime, simply because the opportunity to commit an offence is no longer there (or reduced in practice). Second, an increase in human capital will increase the value of both schooling and work; the former because it will reduce the costs of schooling $c^{ed}(H_0)$ as well as the future value $V(H_1^+)$, the latter because it will increase the current wage as well as the future value $V(H_0)$. So first period crime will decline; whether education will go up is in this context ambiguous.

We can now summarise the way that educational reform and compulsory schooling in particular can affect crime across generations. Meghir and Palme (2005) document the impact of the reform on educational attainment and earnings. For

understanding the impact of the reform they emphasise the importance of distinguishing individuals affected by the reform with parents with just the statutory level of education (60% in that population) to those whose parents had attained a higher level.

An increase in compulsory schooling reduces the available time for crime early on; it increases human capital and thus reduces further the incentive to commit crimes and may increase the chance of remaining in school beyond the new compulsory level. It may also draw increased investments I_1 from parents further increasing H. This reduces crime in the young (school period) ages. As an adult, the result is increased human capital, which as explained before, will reduce *adult* crime. If there is a habit formation aspect of crime (not included above), the early decline will be reinforced. Thus crime will decline relative to the group that was not affected by the reform.

The children of the affected generation are all experiencing the same education system because the reform was rolled out nationally in 1962. They differ by the fact that some have parents who faced the new education system and as a result have more parental education and more resources Y^p. These differences will lead to higher I_0, I_1, and eventually higher educational attainment E_c relative to the children in the comparison group, whose parents did not go through the reform. Educational attainment E_c may increase because, according to mounting evidence, an increase in early investments I_0 improves cognition and social skills and hence reduces the costs of education. In addition the increased resources Y^p allow more transfers to children alleviating somewhat financial constraints for education. These channels imply an increase in both H_0 and H_1 reducing crimes at all life stages, as described above.

For individuals affected by the reform but having parents with more than statutory education the impact is less clear cut. For this group there is no effect on educational attainment. However it changed the way they were educated because it abolished early selection and tracking, which affected primarily this group. It can be argued that quality of education for this group was diluted for this reason *and* because the increase in compulsory schooling, affecting the other group, could have reduced the quality of the peers. For this reason we cannot be confident that H increased for this group. This is why in our empirical analysis we present overall results as well as results separately on the lower socio-economic group.

3 The 1950 Education Reform and the Social Experiment

3.1 The Swedish school system before and after the reform

Prior to the implementation of the comprehensive school reform, pupils attended a common basic compulsory school (*folkskolan*) until grade six. After the sixth grade pupils were selected to either continue one or, in mainly urban areas, two years in the basic compulsory school, or to attend the three year junior secondary school (*realskolan*). The selection of pupils into the two different school tracks was based on their past performance, measured by grades. The pre-reform compulsory school was in most cases administered at the municipality level. The junior secondary school was a prerequisite for the subsequent upper secondary school, which was itself required for higher education.

In 1948 a parliamentary school committee proposed a school reform that im-

plemented a new nine-year compulsory comprehensive school.[4] The comprehensive school reform had three main elements:

1. An extension of the number of years of compulsory schooling to 9 years in the entire country.

2. Abolition of early selection. Although pupils in the comprehensive schools were able to choose between three tracks after the sixth grade - one track including vocational training, a general track, and an academic level preparing for later upper secondary school - they were kept in common schools and classes until the ninth grade.

3. Introduction of a national curriculum. The pre-reform compulsory schools were administrated by municipalities and the pre-reform curriculum varied between municipalities.

3.2 The Social Experiment

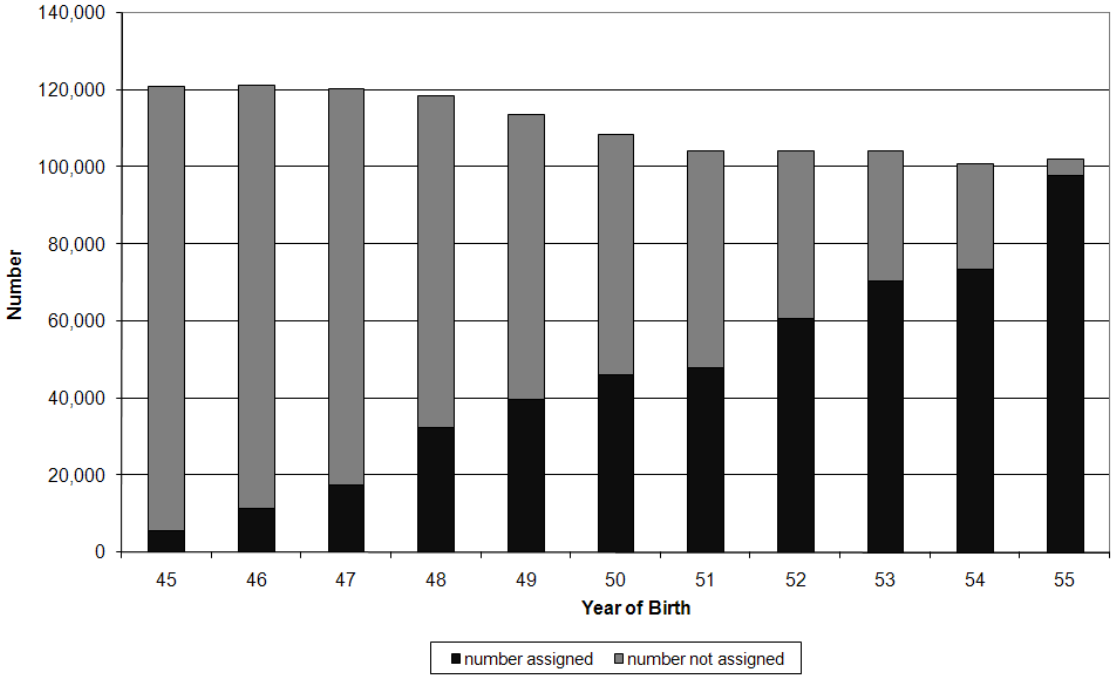
The social experiment with the new comprehensive nine-year compulsory school started during an assessment period between 1949 and 1962, when the final curriculum was decided.[5] The proposed new school system, as described above, was introduced in municipalities or parts of city communities, which in 1952 numbered 1,055 (including 18 city communities). The selection of municipalities was not random. However, the decision to select the areas was based on an attempt to choose areas that were representative for the entire country, both in terms of demographics as well as geographically. At first the National Board of Education contacted the

[4]The school reform and its development are described in Meghir and Palme (2003), Meghir and Palme (2005), and Holmlund (2007). For more detailed reference on the reform, see Marklund (1980) and Marklund (1981).

[5]The official evaluation National School Board (1959) was mainly of administrative nature. Details on this evaluation are also described in Marklund (1981).

municipalities, or sometimes they themselves applied to participate. From this pool of applicants a "representative" sample of municipalities was chosen. Table 12 in the Appendix shows the take up rates of the reform between 1949 and 1962. In our analysis we consider cohorts born between 1945 and 1955. Figure 1 shows the number of observations in each one-year birth cohort and the proportion assigned to the reform.

Figure 1: Proportion of individuals in sample assigned to the reform



4 The Data

4.1 Data Sources and Descriptive Statistics

We use a sample originally obtained from Sweden's population census. To link individuals across generations we used the multi-generation register, provided by Statistics Sweden.[6] We are able to link and use three generations in our analysis:

[6]Statistics Sweden(2003) Flergenerationregistret 2002. En beskrivning av innehåll och kvalitet. Statistics Sweden. Avdelning för Befolknings och Välfärdsstatistik.

the *parent generation* which is the generation directly affected by the reform and it consists of all individuals born in Sweden between 1945 and 1955, their parents the *grandparent generation*, and the *children generation* which are the children of the *parent generation*.[7] This corresponds to 1,340,857 persons, 658,056 males and 655,801 females in the *parent generation*. From the birth certificates we know gender, date and parish of birth. We restrict our sample of the children of the *parent generation*, the *children generation*, to those who have reached the age of criminal responsibility (age 15) in 2008, the last year for which we have crime records. This corresponds to 1,621,758 children, 833,564 sons and 788,194 daughters in the *children generation* that were born between 1959 and 1991.

The reform assignment variable is obtained in two steps. First, we use the name of the church parish of birth in order to obtain the municipality code according to the 1953 Swedish municipality division. Second, based on the year and municipality of birth, we use an algorithm based on historical evidence on reform implementation in each municipality provided by Helena Holmlund and described in Holmlund (2007) to assign reform status to each individual in the sample.

Information on the individual's highest education level was obtained and matched on to our sample from the Swedish National Education Register. For the *grandparent generation* we used data from the 1970 census, which only provides information on individuals younger than age 60 in the year of the census, allowing us to obtain education information for 78.4 percent of mothers and for 65.8 percent of fathers of the *parent generation*. We analyze the effects of the reform separately for those individuals originating from the *low educated grandparent generation*. This is defined as those individuals of the *grandparent generation* with the lowest pre-reform

[7] Even though we have information on biological and adoptive parents and children, we exclude all individuals who have been adopted, or who have adopted children themselves.

statutory level of compulsory schooling. Hence, we analyze the effects of the reform separately for the *parent generation* with low educated fathers, and for the *children generation* with low educated grandfathers, which amounts to roughly 63-65 percent of the sample with available education information.[8]

Information on all convictions in entire Sweden covering the time period between 1981 and 2008 is provided by the Swedish National Council for Crime Prevention (Brå) and has been linked to individuals in our data set. One conviction/court trial often covers several crimes. We have detailed information on the number of crimes the person has been convicted for in each trial, the date of conviction, as well as the penalty for each crime.

Table 1: Number of all convictions in Sweden 1981-2008, for cohorts born 1945-1955 and their children

	Ever convicted		Ever convicted to prison	
	Male	Female	Male	Female
Parent generation				
Number	173,395	46,633	36,870	3,126
Percent	25.31	7.11	5.38	0.48
Low educated fathers				
Number	68,860	16,673	14,525	1,255
Percent	23.96	6.11	5.05	0.46
Children generation				
Number	220,494	69,843	28,588	2,001
Percent	26.45	8.86	3.43	0.25
Low educated paternal grandfathers				
Number	64,976	20,183	7,361	482
Percent	24.68	8.12	2.80	0.19
Low educated maternal grandfathers				
Number	76,887	22,621	9,554	625
Percent	27.78	8.65	3.45	0.24

Notes: We only report the number of convicted children for low educated grandfathers, on both sides of the family.

[8]Table 13 in the data appendix summarizes the number of available observations in each generation and subgroup.

Table 1 shows the number of convicted persons for the two generations, the 1945-1955 cohorts and their children, covered by our data on convictions. Over this time window, 25 percent of all males in the *parent generation* have been convicted at least once, and over 5 percent have been to prison. Only 7 percent of women have been convicted, and 0.5 percent have received a prison sentence. Importantly, the data on criminal convictions only cover the time range between 1981 to 2008, which means that the generations born between 1945 and 1955 will be between the ages of 26 to 63, whereas their children's convictions cover the ages of 15 to 49. The picture for the children generation looks very similar to the one of their parents, with slightly higher percentages of the population having been convicted, possibly attributable to the younger age window. Conviction rates for those with low educated fathers or grandfathers are presented in the lower panel of Table 1. This shows, perhaps surprisingly, that the conviction rates are only marginally higher for this group than for the whole population.

Table 14 in the data appendix shows the crime-age distribution for the entire data set not only covering the cohorts of interest. The largest amount of convictions are for people between 15 and 24, followed by the age range 25 to 34, and further decreasing with age. This pattern of convictions by age is also shown in figures 1 and 2 in the appendix that show the average rate of convictions by age and by cohorts for the cohorts 1970-1989 using men in our children sample.

The stated conviction rates for men of roughly 25 percent is a surprisingly high proportion of the population, which prompted us to look into this in greater detail. First, note that the type of crimes included in our data have to be severe enough to involve a trial and a conviction in court. This includes the more serious traffic violations such as driving without a licence, driving under the influence of alcohol or

drugs, and causing bodily harm, but does not include speeding or parking tickets. As such they do represent serious anti-social behavior. Unfortunately, we were not yet given the specific type of crime for which an individual in our data has been convicted for.[9] However, a good idea of the composition of crime can be obtained in Table 16 in the data appendix where we show a breakdown of type of crime convictions in 2009.

In addition to the data on convictions we have data on all suspected crimes between 1991 and 2009. It includes a variable that gives a detailed code on the type of suspected crime.[10] Although this data overstates actual charges and crimes we use it to provide an idea of the distribution of traffic crimes. Table 17 presents all categories that are related to traffic violations and the number of offences between January 1991 and June 2009. The total number of suspected crimes during this time were 4,073,985 of which 16.9 percent were traffic crimes. Again, all of these traffic crime categories are severe violations. Additional support of such high conviction rates in Sweden is provided by previous studies that have shown similar conviction rates, see Hjalmarsson and Lindquist (2010a) and Hjalmarsson and Lindquist (2010b).

4.2 The Association of Education, Family and Criminal Behavior

Table 2 shows the Linear Probability Model estimation results of whether an individual has ever been convicted (column (1)) and ever been convicted to a prison sentence (column (2)) on years of own schooling as well as on years of father's and

[9] We are waiting to obtain a variable that indicates the type of crime from the Brå crime registry.

[10] Detailed coding of crime types in: Kodning av brott, Anvisningar och regler, Version 8.0, Reviderad 1. Juli 2010, brå brottsförebyggande rådet.

mother's schooling for the children generation. The estimations for the sample of men born in Sweden between 1945 and 1955 are presented in the upper panel A, and the ones for the sons in the lower panel B. All regression models include both a full set of dummy variables for the birth cohorts and birth municipalities to control for area and cohort fixed effects.[11] Column (1) in panel A shows that one year of own schooling for men in the parent generation is associated with a decrease of the probability of a conviction by 1.9 percentage points, which translates to a reduction in the conviction rate of 7.5 percent. One year of schooling for men decreases the probability of a prison sentence by 0.8 percentage points, or a reduction in the total share of men convicted to prison by 15 percent, see column (2). Panel B includes own schooling and both parents' education for sons, and hence estimates the effect of an increase of schooling of one parent controlling for assortative mating effects. One year of own schooling is associated with a 2.64 (0.59) percentage points decrease in the probability of sons' conviction (imprisonment), corresponding to a 9.96 (17.06) percent reduction in the share of convicted (imprisoned) sons. A years decrease in father's schooling is associated with an increase in the son's conviction rate (prison rate) of 0.568 (0.118) percentage points or equivalently a 2.15 (3.44) percent reduction in the share of convicted (imprisoned) sons. The equivalent figure for a decrease in mother's education by a year is a 2.261 (4.198) percent decrease in the share of convicted (imprisoned) sons.

We present the relationship between crime and the levels of education in the data appendix (see Tables 18 and 19), revealing a steep decline in crime participation associated with higher levels of own and parental education. A similar decline is also recorded for incarceration rates.

[11] When computing the standard errors we cluster by birth municipality, which is conservative and makes the standard errors comparable to the ones we estimate for the effect of the reform below.

Table 3 shows the results of a Probit model of the association between having a father who has been convicted (or to prison) and the father's age of birth on the probability of his son ever been convicted (or to prison). The probability of ever been convicted increases by over 15 percentage points if a son has a convicted father. This corresponds to a 63 percent increase of the total share of convicted sons. If he had a father who had been convicted to a prison sentence the probability of an own prison sentence increases by over 7 percentage points, which translates to a 134 percent increase in the share of convicted sons.

Overall, there is a large and significant negative correlation between own education and own criminal participation, both for the probability of a conviction and the probability of a conviction to prison. This holds for two subsequent generations. Furthermore, our evidence strongly supports the notion that higher education of both mothers and fathers is associated with a lower probability that their sons will be engaged in crime. Finally, we find a strong and highly significant positive intergenerational link between father's and son's criminal behavior.

Table 2: Linear probability model estimates of the association between own or father's education and the probability of being convicted and the probability of being convicted to a prison sentence. Men born between 1945-1955 and their sons.

Panel A	Men born 45-55	
Dependent variables	Probability conviction $\bar{p} = 0.2531$	Probability prison $\bar{p} = 0.0538$
Years of schooling, own	-1.916***	-0.806***
	(0.093)	(0.057)
Corresponding percentage change	-7.570	-14.981
Birth cohort/municipality dummies	y	y
Observations	662,875	662,875
Panel B	**Sons of men born 45-55**	
Dependent variables	Probability conviction $\bar{p} = 0.2645$	Probability prison $\bar{p} = 0.0343$
Years of schooling, own	-2.635***	-0.585***
	(0.029)	(0.013)
Corresponding percentage change	-9.962	-17.055
Years of schooling, father	-0.568***	-0.118***
	(0.035)	(0.014)
Corresponding percentage change	-2.147	-3.440
Years of schooling, mother	-0.598***	-0.144***
	(0.026)	(0.009)
Corresponding percentage change	-2.261	-4.198
Birth cohort/municipality dummies	y	y
Observations	675,625	675,625

Notes: Results are scaled by 100. Level of observation in upper panel are men born in Sweden between 1945 and 1955. Level of observation in lower panel are sons of parents born between 1945 and 1955. Robust standard errors in parentheses, clustered by birth municipality. All regressions in the upper panel include a full set of birth cohort dummies and birth municipality dummies. In the lower panel a full set of municipality dummies and birth cohort dummies of the child was included. Significance levels *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Probit model marginal effects of the association between the probability of ever been convicted and the father having ever been convicted. Marginal effects are scaled by 100. Estimations for different education levels of grandfathers. Sons of men born in Sweden between 1945 and 1955.

Panel A	Sons of men born 45-55	
Dependent variables	Probability conviction $\bar{p} = 0.245$	Probability prison $\bar{p} = 0.054$
Father convicted/Father prison	15.457***	7.248***
	(0.260)	(0.183)
Percent change	63.090	134.222
Age of father at birth of son	-1.759***	-0.323 ***
	(0.019)	(0.007)
Observations	559,085	559,085
Panel B	Sons of men born 45-55 with low educated father	
Dependent variables	Probability conviction $\bar{p} = 0.245$	Probability prison $\bar{p} = 0.021$
Father convicted/Father prison	14.755***	7.116***
	(0.245)	(0.243)
Percent change	60.224	338.857
Age of father at birth of son	-1.765***	-0.315***
	(0.022)	(0.007)
Observations	241,716	241,716

Notes: Level of observation in upper panel are sons of men born in Sweden between 1945 and 1955, and in the lower panel sons of men born between 1945 and 1955 who have a low educated father. Robust standard errors in parentheses, clustered by birth municipality of sons. All regressions include a full set of birth cohort dummies of fathers. Significance levels *** p<0.01, ** p<0.05, * p<0.1.

5 Identification Strategy

The main outcome variable we use is whether an individual was ever convicted during the observation window 1981-2008. We also consider whether someone has ever received a prison sentence. Finally, we consider the number of convictions.

All the analysis is done for males only and we distinguish them by the education of the *grandparent generation*. We present two sets of estimates. The first relate to the impact of the reform on the parent generation, i.e. the generation affected by the educational reform directly. The second relate to the impact of the reform on the children of the parent generation. Referring back to the discussion on how the reform affects crime, for the parent generation the estimated impact related to adult crime only because the youngest person in the sample is 26 when the crime records start. Hence, the effect is not attributable to the simply keeping the kids off the streets by getting them to attend school. For the child generation the impact is attributable to better parenting and increased investments in children; it is not due to differences in the schooling system they are facing, which is the same for the treatment and comparison group children.

If the reform had been randomised we could just compare crime rates across treatment and control municipalities at each point in time. However such randomisation did not take place and as a result we need further identifying assumptions so as to account for potential differences across treatment and control municipalities. Assuming that the propensity to commit a crime can be written as

$$y^*_{i,m,t} = F(\epsilon, t),$$

where t denotes cohort and ϵ denotes unobserved heterogeneity. As in Athey and

Imbens (2006) and Altonji and Blank (1999) the key assumptions are that (i) the distribution of ϵ, while differing across municipalities, does not change from cohort to cohort and (ii) F is strictly monotonic in ϵ. These "difference-in-differences" assumptions are sufficient to identify the impact of the reform, but do imply restrictions on the model for crime participation equation 2.3. In particular it requires that if there is any unobserved heterogeneity, whose distribution depends on the municipality, this has to be summarised by a single unobserved factor, say entering through human capital, and whose distribution is the same across cohorts; the propensity to commit a crime must be *strictly* monotonically related to this unobserved factor. In practice this assumes that increases in human capital always reduce crime and that this is never reversed by the reduction in the probability of being caught, which does not seem to be a very strong assumption.

With a discrete outcome variable, such as ours we need additional assumptions if we are to obtain a point estimate.[12] We thus assume that heterogeneity and cohort effects are additively separable and we use both the linear probability model and the logit model as alternative distributional assumptions. Specifically we assume

$$y^*_{i,m,t} = \alpha + \beta_1 R_{i,m,t} + \gamma_1' t_i + \gamma_2' M_i + \epsilon_{i,m,t},$$

where $y^*_{i,m,t}$ is the latent crime "intensity" outcome observed for person i born in municipality m and in birth cohort t. A conviction corresponds to $y^*_{i,m,t} > 0$. $R_{i,m,t}$ is the reform indicator, which equals one if individual i belongs to a municipality and cohort that has been assigned to the new school system; t_i is a vector of dummy variables indicating to which cohort individual i belongs to and M_i is a vector of dummy variables indicating in which municipality individual i was born.

[12] This is the case with all papers that employ difference-in-differences analysis of discrete outcomes (see Athey and Imbens (2006)).

Identification follows by assuming that

$$E(\epsilon_{i,m,t}|R_{i,m,t}, t_i, M_i) = 0.$$

The key assumptions are that the distribution of the unobservables are invariant over time and that the crime trends in the treated groups are the same as those in the comparison groups. The long time series of the data where we observe criminal activity later in life of cohorts who were not affected by the reform but went to school in municipalities that later implemented the reform and indeed also for several cohorts that went to the new school system allows us to present some corroborative evidence in support of the common trends assumption.

Our first approach is the linear probability model where

$$\Pr(y^*_{i,m,t} > 0|R_{i,m,t}, t_i, M_i) = \alpha + \beta_1 R_{i,m,t} + \gamma_1 t_i + \gamma_2 M_i.$$

This is straightforward to estimate by GLS. However, the linear probability model has the well known disadvantage that it does not restrict the probabilities to lie on the [0,1] interval. The main reason we use it is for computational convenience, given the fact that there are about 1,000 municipality and 11 cohort fixed effects.

As an alternative, we also estimate a logit model by minimum distance: first we group the data by municipality and cohort and estimate the within-cell conviction probability (P_{mt}). We then use minimum distance to impose the restriction that this probability is generated from a logistic distribution with a linear index as in the latent equation above by fitting the log-odds ratio as follows

$$\log(\frac{P_{mt}}{1-P_{mt}}) = \delta_0 + \delta_1 R_{m,t} + \delta_2' t_i + \delta_4' M_i.$$

We drop all cells where the log odds ratio is not defined, which amounts to about 6% of cells. Implicitly the logit and the LP models deal with such cells as well as with the nonlinear form of the probabilities in a different way and hence it is interesting to see if we get different results with the two approaches.

6 Results

6.1 The First Generation

Tables 4 through 7 show the estimates of the causal effect on criminal behavior of being assigned to the reform. Tables 4 and 5 show the results for the probability of having ever been convicted for the linear probability and logit model, respectively. Table 6 shows the results for the probability of having ever been convicted to a prison sentence using the linear probability model and, finally, Table 7 shows the effect of the reform on the intensity to engage in crime measures by the number of crimes an individual was convicted for. Throughout, we present results for the entire sample of men as well as separately for those coming from homes with low educated fathers. All specifications include fixed effects for birth municipality as well as birth cohorts and the standard errors are corrected for clustering within municipality of birth, allowing for both spatial and serial correlation.[13]

To avid the computational difficulties involved in estimating a logit model with 1,000 municipality fixed effects and 11 cohorts we use a minimum distance procedure.

[13] On clustering the standard errors see Moulton (1990).

We collapse the sample to 10,744 municipality-cohort cells by computing the log-odds ratio within each cell. For 691 municipality-cohort cells the proportion of observed crime was zero and hence the log-odds ratio is not defined. For 108 cells we cannot assign the reform status, which leaves us with 9,949 municipality-cohort observations. We then regress the log-odds ratio on the municipality and cohort dummies as well as on the reform indicator using GLS. Each cell was weighted by $\sqrt{p_c(1-p_c)N_c}$, where N_c is the cell size and p_c is the within cell probability of a conviction. The corresponding marginal effects for different cohorts of the logit model are presented in Table 5. For the estimates when prison sentences are used as dependent variable this is not possible since the proportion of prison sentences is too small, which prevents us from computing the log-odds ratios.

Column (1) in Tables 4 and 5shows a negative, but insignificant effect of the reform on the probability of having ever been convicted for the entire sample born between 1945 and 1955 and, in the lower panel of the tables, a similar result for those originating from homes with low educated fathers. However, these results may be diluted by the fact that the oldest birth cohorts in our data are quite old when we start to collect data on convictions and therefore may have ended their criminal careers. For example, the oldest cohort included, those born in 1945, are aged 36 when we start to record their criminal behavior. This is an age by which a large proportion have already stopped committing crimes.

To account for this possibility, we also present results where we have restricted the sample to younger birth cohorts. Columns (2) through (6) in Tables 4 and 5 show results for the impact of the reform on different cohort groups. Although, we clearly loose some precision as we reduce the number of cohorts considered, the results show a clear pattern for both the entire sample and those originating from

homes with low educated fathers: for younger cohorts, which are observed at the age range when more crimes are committed, the reform seems to have increasingly strong effects. The estimate for the youngest cohort, born in 1954 or 1955, is highly significant suggesting a 1.3 percentage points decrease in crime; this corresponds to a 5 percent, decrease in the probability of ever been convicted as a result of being assigned to the post reform school system. The effect is somewhat larger in the subsample of men born in homes with low educated fathers. These results are consistent with findings by Lochner and Moretti (2004) and Machin, Marie, and Vujić (2010) for the US and the UK respectively, who show that increases in compulsory schooling led to reductions in crime.

To analyze if the education reform also decreased the amount of crimes individuals were convicted for we use an alternative outcome variable that measures the total number of crimes an individual has been convicted for between 1981 and 2008. The average number of convicted crimes for men in the first generation ranges from 1.2 to 1.8, depending on cohort and father's education level. This outcome variable is characterized by a large mass point at zero and overdispersion with a variance roughly ten times larger than the mean. To estimate the model on this outcome variable we use a negative binomial model. Table 7 presents the marginal effects from these estimates. The estimates are negative and significant at the 10 percent level for those born between 1950 and 1955 in homes with low educated fathers. The effect seems to be somewhat stronger for those born between 1952-1955 and 1953-1955, respectively. These results are in line with the findings on the probability of having ever been convicted.

Table 4: Linear Probability Model estimates of the effects of the reform on the probability of ever been convicted. Marginal effects are scaled by 100. Estimations by birth cohorts for the entire sample and for men with low educated fathers separately. Men born in Sweden between 1945 and 1955.

Panel A	Men born 45-55					
Cohorts	45-55	50-55	51-55	52-55	53-55	54-55
Probability conviction	0.2531	0.2683	0.2714	0.2746	0.2783	0.2837
Reform	-0.645	-0.456	-0.532*	-1.028***	-1.076**	-1.329***
	(0.405)	(0.305)	(0.318)	(0.396)	(0.490)	(0.479)
Percentage change	-2.548	-1.700	-1.960	-3.744	-3.866	-4.685
Share convicted						
Observations	622,583	319,093	263,592	210,399	157,155	103,761
Panel B	Men born 45-55 with low educated father					
Cohorts	45-55	50-55	51-55	52-55	53-55	54-55
Probability conviction	0.2396	0.2517	0.2552	0.2588	0.2631	0.2693
Reform	-0.263	-0.494	-0.579	-1.119*	-1.103	-2.094**
	(0.304)	(0.434)	(0.486)	(0.661)	(0.912)	(0.970)
Percentage change	-1.098	-1.963	-2.269	-4.324	-4.192	-7.776
Share convicted						
Observations	264,679	150,620	125,952	101,266	76,207	50,222

Notes: The table presents results from a linear probability (LP) estimation with weighted least squares. The used weights are: $\sqrt{x'b(1-x'b)}$, which are obtained from a first stage OLS estimation. The level of observation are men born between between 1945-1955, or men born in cohorts indicated at the top of each column; Reform is a dummy variable indicating the reform status of the index person. All regressions include a full set of birth municipality dummies and birth cohort dummies. Robust Standard errors in parentheses, clustered by municipality of birth. *** p<0.01, ** p<0.05, * p<0.1.

6.2 The Child Generation

Tables 8 and 9 report the results of the difference-in-differences estimation of the effects of the school reform on criminal participation of the child generation for the probability of ever been convicted and prison convictions, respectively. Again, for the probability of being convicted, we additionally estimate a logit model based on cohort-municipality cells. We estimate two specifications. In the first one, we estimate the effects of a father who has been subjected to the new school system on

Table 5: Logit model marginal effects of the reform on the probability of ever been convicted. Marginal effects are scaled by 100. Estimations by birth cohorts for the entire sample and for men with low educated fathers separately. Men born in Sweden between 1945 and 1955.

Panel A	Men born 45-55					
Cohorts	45-55	50-55	51-55	52-55	53-55	54-55
Probability conviction	0.253	0.268	0.271	0.275	0.278	0.284
Reform	-0.023	-0.284	-0.434	-1.274*	-0.999	-1.395
	(0.281)	(0.422)	(0.500)	(0.654)	(0.911)	(1.361)
Percentage change	-0.091	-1.057	-1.598	-4.639	-3.591	-4.916
Share convicted						
Observations	622,583	319,093	263,592	210,399	157,155	103,761
Panel B	Men born 45-55 with low educated father					
Cohorts	45-55	50-55	51-55	52-55	53-55	54-55
Probability conviction	0.234	0.252	0.255	0.259	0.263	0.269
Reform	-0.022	-0.272	-0.417	-1.227*	-0.965	-1.351
	(0.271)	(0.405)	(0.480)	(0.630)	(0.879)	(1.318)
Percentage change	-0.093	-1.081	-1.633	-4.741	-3.666	-5.015
Share convicted						
Observations	264,679	150,620	125,952	101,266	76,207	50,222

Notes: We only report marginal effects of the logit model; Reform is a dummy variable indicating the reform status of the index person. All regressions include a full set of birth municipality dummies and birth cohort dummies. The level of observation are men born between 1945-1955, or men born in cohorts indicated at the top of each column. Robust Standard errors in parentheses, clustered by municipality of birth of the individual. *** p<0.01, ** p<0.05, * p<0.1.

child criminal behavior and in the second one the corresponding effects of a mother going through the post reform school system.

We distinguish between four samples. The first is the sample of individuals whose father was born between 1945 and 1955 for which we estimate the effect of their father's school reform assignment on their probability of a conviction. The second sample are individuals whose mother was born between 1945 and 1955, for which we estimate the effect of their mother's school reform assignment on their son's criminal behavior. The other two samples are individuals originating from the *low educated*

Table 6: Linear Probability Model estimates of the effects of the reform on the probability of ever been convicted to a prison sentence. Marginal effects are scaled by 100. Estimations by birth cohorts for the entire sample and for men with low educated fathers separately. Men born in Sweden between 1945 and 1955.

Sample	Men born 45-55	Men born 45-55 with low educated fathers
Probability prison	$\bar{p} = 0.0538$	$\bar{p} = 0.0505$
Reform	-0.149	-0.049
	(0.160)	(0.170)
Corresponding percentage change share prison	-2.770	-0.974
Observations	622,583	264,679

Notes: The table presents results from a linear probability (LP) estimation with weighted least squares. The used weights are: $\sqrt{x'b(1-x'b)}$, which are obtained from a first stage OLS estimation. The level of observation are men born between 1945-1955. 5.38% of all men in the considered sample have been convicted to prison at least once. 5.05% of the considered sample who have a low educated father have been to prison. Robust Standard errors in parentheses, clustered by municipality of birth. *** p<0.01, ** p<0.05, * p<0.1. Reform is a dummy indicating the reform status of the index person. All regressions include a full set of birth municipality dummies and birth cohort dummies.

grandparent generation. More specifically, the third sample is defined like the first one, but restricted to those individuals whose grandfather on the father's side is low educated. Finally, the fourth sample is restricted to those individuals within the second sample whose grandfathers on their mother's side are low educated.

When we collapse the data set by cohort-municipality level as a first step to estimating the logit model, we do it by father's cohort-municipality level which gives us 10,607 cells for the father's sample, and 10,247 for the sample with low educated paternal grandfathers. For the specification with mother's reform assignment we collapse the data by mother's cohort-municipality level which leads to 10,647 for the entire sample and 10,324 for the low educated maternal grandfathers sample.

Table 8 reports the marginal effects of the linear probability model for a conviction (denoted by LP), the log-odds ratio estimate for the logit model, and the corresponding marginal effect of the logit model, both for the sample of all sons, and for sons whose grandfather was low educated. It can be seen that the reform

Table 7: Marginal effects of negative binomial estimation of the effects of the reform on the number of crimes men have been convicted for between 1981-2008. Estimations by birth cohorts for the entire sample and for men with low educated fathers separately. Men born in Sweden between 1945 and 1955.

Sample	Men born 45-55					
Cohorts	45-55	50-55	51-55	52-55	53-55	54-55
$\bar{y}$	1.309	1.578	1.646	1.696	1.748	1.828
Reform	0.027	-0.071	-0.038	-0.122	-0.078	-0.121
	(0.035)	(0.062)	(0.067)	(0.086)	(0.083)	(0.092)
Observations	622,583	319,093	263,592	210,399	157,155	103,761
Sample	Men born 45-55 with low educated fathers					
Cohorts	45-55	50-55	51-55	52-55	53-55	54-55
$\bar{y}$	1.253	1.473	1.539	1.587	1.647	1.717
Reform	-0.030	-0.137*	-0.117	-0.250**	-0.236*	-0.177
	(0.047)	(0.081)	(0.096)	(0.124)	(0.128)	(0.168)
Observations	264,679	150,620	125,952	101,266	76,207	50,222

Notes: Margnial effects of the negative binomial estimation are reported. Reform is a dummy indicating the reform status of the individual. Robust Standard errors in parentheses, clustered by individual's municipality of birth. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a full set of birth municipality dummies and birth cohort dummies of the index person. The estimation method is a negative binomial model to relax the Poisson model assumption of equal conditional mean and conditional variance of the outcome variable, but maintaining an exponential mean.

significantly reduces the probability of having ever been convicted for the sons of fathers who went through the reform by 0.6 percentage points, see column (LP). Since the average share of convicted individuals in this cohort was about 26.5 percent, reduction in criminality was approximately 2.5 percent.

The log-odds-ratio estimates from the logit model translate into a marginal effect of a 0.646 percentage points decrease in the probability of a conviction, which is very similar to the marginal effect obtained from the linear probability model. Hence, the logit model suggests similar to the linear probability model, that father's reform assignment significantly reduces the total share of convicted men by about 2.5 percent.

Consider now the sample of sons with a low educated paternal grandfather. The impact of their father being assigned to the reform is a reduction of the probability of conviction by 1.02 percentage points. This translates into a 4.13 percent decrease in the share of convicted sons of low educated grandfathers (24.7 percent). The marginal effect of the logit model is of the same magnitude (-0.97 percentage points, and -3.94 percent decrease in total share of convicted).

There are no significant effects of reform assignment of mothers on the probability of a conviction of their sons neither for the entire sample of sons, nor for those with a low educated maternal grandfather. Also, as can be seen in Table 9, the results of the linear probability model for the effects of the reform status of father and mother on the probability of a prison sentence of sons show no significant effects.

Table 10 shows the negative binomial estimates of the marginal effects for the intensive margin on the number of crimes convicted for between 1981 and 2008 for the children generation. The top panel A shows the results based on father's reform status for all children and those with low educated grandfathers, respectively. The bottom panel B shows the corresponding results based on mother's reform status. The estimate based on father's reform status shows a significant negative effect of the reform based on father's reform status, which is in line with our findings on the probability of being convicted.

We repeated the analysis for the children generation using the suspected crime data mentioned above. Descriptive tables on suspected crime rates for the relevant sample of children is presented in the appendix in table 15. Interestingly, the rates of ever been suspected for a crime including and excluding all traffic crimes are very similar. We find a negative but not significant relationship between father's reform assignment and sons probability of ever been suspected for a crime.[14]

[14]The results are provided by the authors upon request.

Table 8: Linear Probability Model estimates and Logit Model estimates of the marginal effects of father's and mother's reform assignment on the probability of their sons having ever been convicted. Marginal effects are scaled by 100. Estimations for the entire sample and for sons with low educated grandfathers separately. Sons of men and women born in Sweden between 1945 and 1955.

Panel A: Father's reform assignment						
Sample	Sons of men born 45-55			Sons of men born 45-55 with low educated fathers		
Probability conviction	$p = 0.265$			$p = 0.247$		
Specification	LP	Logit	Logit Marginal effect	LP	Logit	Logit Marginal effect
Reform_father	-0.650***	-0.033**	-0.646***	-1.02***	-0.052***	-0.972***
	(0.219)	(0.015)	(0.296)	(0.361)	(0.020)	(0.366)
Corresponding percentage Change share convicted	-2.456		-2.443	-4.129		-3.938
Observations	563,754	563,754	563,754	243,082	243,082	243,082
Panel B: Mother's reform assignment						
Sample	Sons of women born 45-55			Sons of women born 45-55 with low educated fathers		
Probability conviction	$\bar{p} = 0.265$			$\bar{p} = 0.278$		
Specification	LP	Logit	Logit Marginal effect	LP	Logit	Logit Marginal effect
Reform_mother	-0.159	0.012	0.225	-0.041	0.021	0.419
	(0.249)	(0.014)	(0.277)	(0.331)	(0.018)	(0.363)
Corresponding percentage Change share convicted	-0.600		0.851	-0.147		1.509
Observations	595,138	595,138	595,138	255,075	255,075	255,075

Notes: The top panel shows the estimated effects of father's reform status Reform_father on sons crime, and the bottom panel shows the estimated effects of mother's reform status Reform_mother on son's crime. 26.45% of men in the considered sample have been convicted at least once. Of those men who have a low educated grandfather on their father's side 24.68% have been convicted. Of those men who have a low educated grandfather on their mother's side 27.78% have been convicted. Columns (1) and (4) present linear probability model (LP) estimation with weighted least squares. The used weights are: $\sqrt{x'b(1-x'b)}$, which are obtained from a first stage OLS estimation. The level of observation are sons born before 1994, whose parents were born between 1945-1955. Column (2) and (5) show the results of a logit estimation; we report the coefficient of the log odds ratio coefficient. The marginal effect for the logit model are presented in columns (3) and (6). We reduced the sample on father's or mother's birth municipality and birth cohort level to estimate the logit model. Reducing the sample to the following parent's cohort-birth municipality cells: for sons with a father born 1945-1955 to 10,607 cells, for sons whose father has a low educated father to 10,247 cells, for sons whose mother was born 1945-1955 to 10,647 cells, and whose mother had a low educated father to 10,324 cells. Robust Standard errors in parentheses, clustered by municipality of birth. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a full set of birth municipality dummies and birth cohort dummies of father's or mother's.

Table 9: Linear Probability Model estimates of the effects of father's and mother's reform assignment on the probability of their sons having ever been convicted to a prison sentence. Marginal effects are scaled by 100. Estimations for the entire sample and for sons with low educated grandfathers separately. Sons of men and women born in Sweden between 1945 and 1955.

Panel A: Father's reform assignment		
Sample	All Sons	Sons with low educated grandfather
Probability prison	$p = 0.034$	$p = 0.028$
Reform_father	0.015	0.005
	(0.093)	(0.123)
Corresponding percentage Change share prison	0.448	0.169
Observations	563,754	243,082
Panel B: Mother's reform assignment		
Sample	All Sons	Sons with low educated grandfather
Probability prison	$p = 0.034$	$p = 0.028$
Reform_mother	0.038	0.004
	(0.093)	(0.130)
Corresponding percentage change share prison	1.093	0.108
Observations	595,138	255,075

Notes: The reported estimates are based on a weighted least squares estimation. 3.43% of all men in the sample have been to prison. 2.80% of men who have a low educated grandfather on their father's side have been convicted to prison. 3.45% of men who have a low educated grandfather on their mother's side have been convicted to prison. All regressions include either a full set of birth municipality dummies of father or mother, and birth cohort dummies of father or mother. Robust Standard errors in parentheses, clustered by municipality of birth of mother or father. *** p<0.01, ** p<0.05, * p<0.1.

A possible mechanism for the inverse effect of the reform on the criminal behavior of the child generation is through fertility and family formation. An implication of both the well known quality-quantity model and the time allocation model is that increased educational attainments and earnings prospects will decrease the demand for having children. If there is a trade off between the number of children and parental investments in each child, it may make up an inverse relation between

Table 10: Marginal Effects of negative binomial estimation of father's or mother's reform assignment on the number of crimes sons are convicted of between 1981-2008. Estimations for the entire sample and for sons with low educated grandfathers separately. Sons of men and women born in Sweden between 1945 and 1955.

Panel A: Father's reform assignment		
Sample	All Sons	Sons with low educated grandfather
Average number convicted crimes	1.245	1.210
Reform_father	-0.065**	-0.040
	(0.031)	(0.050)
Observations	563,754	243,082
Panel B: Mother's reform assignment		
Sample	All Sons	Sons with low educated grandfather
Average number convicted crimes	1.538	1.442
Reform_mother	0.010	0.069
	(0.046)	(0.055)
Observations	595,138	255,075

Notes: Marginal effects are reported, $\bar{y}$ is the average number of convicted crimes in the sample. Experiment_father or Experiment_mother is a dummy indicating the reform status of the father or mother of the son. Robust Standard errors in parentheses, clustered by municipality of birth of the son's father (Panel A) or mother (Panel B). *** p<0.01, ** p<0.05, * p<0.1. All regressions include a full set of birth municipality dummies and birth cohort dummies of father (Panel A) or mother (Panel B). The estimation method is a negative binomial model to relax the Poisson model assumption of equal conditional mean and conditional variance of the outcome variable, but maintaining an exponential mean.

being assigned to the reform and the inverse relation to child criminality.

To further investigate this possible mechanism we estimate the effect of the reform on fertility behavior. We use four different outcomes: the total number of children, the probability of having children, the age at birth of first child and, finally, the probability of having children while still a teenager. Since we only find significant effects of the reform on sons to fathers affected by the reform, we restrict the sample to men.[15] As can be seen in the results of these estimates shown in table 11, the only significant effect is on the probability of having a child as a teenager.

[15]The results for women are qualitatively the same, the only difference is that women who were assigned to the reform are less likely to ever have a child on a 10 percent significance level.

However, since only 1.3 percent of the population have children as teenagers and, more importantly, the reform only decreased this probability by between 0.2 and 0.3 percent, we conclude that we cannot attribute the effect of the reform to changes in fertility behavior. The results on all other fertility outcome variables shown are estimated to be zero.

Table 11: Estimation of the effects of the reform on the probability of ever having a child, the number of children, the age at birth of first child and the probability of teenage paternity. Marginal effects are scaled by 100. Estimations by birth cohorts for the entire sample and for men with low educated fathers separately. Men born in Sweden between 1945 and 1955.

Dependent variable	ever child	number children	age birth first child	teenage
Specification	LP	Poisson	Neg binomial	LP
Sample:		Men born 45-55		
Average dep var	0.813	1.896	27.054	0.017
Reform	-0.093	-0.004	0.106	-0.263**
	(0.185)	(0.007)	(0.075)	(0.106)
Observations	622,583	622,583	505,679	622,583
Sample:		Men born 45-55 with low educated fathers		
Average dep var	0.822	1.912	26.524	0.019
Reform	-0.096	0.001	0.064	-0.210**
	(0.273)	(0.010)	(0.048)	(0.100)
Observations	264,679	264,679	217,517	264,679

Notes: Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered by municipality level. All estimations include a set of municipality dummies and cohort dummies. The level of observation are men born between 1945 and 1955, all men born between 1945-1955. Experiment is a indicator variable for reform assignment.

6.3 The common trends assumption

As is well known, an identifying assumption underlying the differences-in-differences estimator is that any trend in the outcome variable is common in the treatment and comparison groups over the period of comparison. This assumption is untestable

because it relates to the counterfactual change in the treatment group. However, an indication can be obtained by testing whether the trends are common in the two sets of groups before the reform and indeed after the reform as well.

In our sample we have 12 groups of municipalities indexed by which cohort was first assigned to the reform. To implement the common trends test we used only the municipalities that first implemented the reform for the 1947 cohort onwards (i.e. 10 of the 12 set of municipalities). We compared the trend of criminal behavior of individuals across these municipalities for all cohorts that were not affected by the reform. The pooled regression of these groups is $y = \alpha + \beta t + \gamma' m + \delta' t * m + \nu$, where m is a set of dummies indicating the group to which the municipality belongs based on the cohort for which it first implemented the reform; t is a linear trend that represents the cohorts 1945-1955. A joint test of $\delta_1 = \delta_2 = ... = \delta_{10} = 0$ gives a F statistic of $F(9;7{,}090)=1.15$ with $P=0.323$, with 7,090 cohort-municipality cells before treatment. This implies the hypothesis of common trends in crime for the pre-treatment cohorts for all groups of municipalities cannot be rejected.

For post-treatment trends in crime we only compare crime between the municipalities that implemented the reform for cohorts born before 1954. We compare the criminal behavior of individuals across these municipalities for the cohorts affected by the reform. A joint test of equality of the coefficients on the interaction term of the above pooled model yields F(9, 4,808) with P=0.1303, where 4,808 is the number of cohort-municipality cells that are treated. This means that the hypothesis of a common trend in criminal behavior for the treated cohorts is the same across the groups of municipalities that implemented the reform for different cohorts cannot be rejected. Both these tests are strong evidence in favor of the key identifying assumption for our difference-in-differences approach to the problem.

7 Conclusion

We find a very strong relation between own educational attainments and criminal behavior in our data: every year of additional education is associated with a decrease of about 7.5 percent in the probability of being convicted for a crime and a 15 percent decrease in the probability of a prison sentence. We also find a strong relation between parental education and crime. On average, one year of additional education of the father is associated with an about 2.1 percent decrease in the probability of being convicted and an about 3.5 percent decrease in the probability of a prison sentence. To what extent do these strong correlations give guidance to the likely effects of a policy intervention designed to improve educational attainment of the least skilled? The staggered implementation of a major Swedish education reform in the 1950s and early 60s, which among other things increased compulsory schooling, provides an ideal opportunity to answer this question.

Using administrative data that compares children of the same cohorts, but educated under different systems, we find strong negative and significant effects of the reform, particularly when we focus on cohorts whose criminal activity is recorded at a younger age. Thus, for the youngest cohorts, born between 1954 and 1955, the point estimate suggests a 1.3 percentage points, corresponding to 5 percent, decrease in the probability of being convicted from being assigned to the post reform school system. In the group from homes with low educated fathers the effect seems to be somewhat larger, which is consistent with a larger effect on both educational attainments and labor earnings in this group found in Meghir and Palme (2005).

The striking result of this paper however is the effect of the reform on the children of those originally affected: there is a significant effect of paternal assignment to the reform on the probability of being convicted corresponding to an average reduction of

about 2.5 percent. However, we found no significant effect from mother's assignment to the reform. The intergenerational effect of education on criminal behavior can operate through several different channels. The result that the reform had a stronger effect on criminal behavior in the child generation - together with the result obtained in Holmlund, Lindhal, and Plug (forthcoming) that the reform had a weak, although significant, effect on the educational attainments in the next generation - suggests that effect did not operate solely through increased education in the child generation. Another conceivable channel is through changes in fertility induced by the reform. Our data allowed us to estimate reform effects on four different outcomes related to fertility behavior. Since our results suggest that the effect operated through father's reform assignment, we restricted the sample to men. The only significant effect was a decreased probability of having a child as teenager. However, the effect was far too small to serve as a major explanation of the decreased criminality in the next generation.

Given the data at our disposal, there are several channels that we are not able to distinguish between. The most obvious one is that improved earnings possibilities provide the parents with improved economic resources to invest in the upbringing of their children. However, since crime is also a manifestation of a basic lack of social skills or other emotional deficits it is also conceivable that the effect of parental education works through other channels than solely economic resources. A possible interpretation of the fact that we find significant effects of the father's rather than the mother's reform assignment on the son's criminal behavior is consistent with the notion that these alternative explanations may be important, in view of the fact that female labor force participation is very high in Sweden.

References

Altonji, J. G., and R. M. Blank (1999): "Chapter 48 Race and gender in the labor market," vol. 3, Part 3 of *Handbook of Labor Economics*, pp. 3143 – 3259. Elsevier.

Athey, S., and G. W. Imbens (2006): "Identification and Inference in Nonlinear Difference-In-Differences Models," *Econometrica*, 74(2), 431–497.

Becker, G. (1981): *A treatise on the family*. Harvard University Press, New York.

Björklund, A., M. Lindahl, and E. Plug (2006): "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *Quarterly Journal of Economics*, 121(3), 999–1028.

Carneiro, P., C. Meghir, and M. Parey (2007): "Maternal education, home environments and the development of children and adolescents," *IZA Discussion paper No. 3072*.

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

Freeman, R. B. (1996): "Why Do So Many Young American Men Commit Crimes and What Might We Do About It?," *The Journal of Economic Perspectives*, 10(1), pp. 25–42.

——— (1999): "The Economics of Crime," vol. 3, Part 3, Chapter 52 of *Handbook of Labor Economics*. Elsevier.

Gallipoli, G., and G. Fella (2006): "Education and Crime over the Life Cycle," *Working Paper*.

Gould, E. D., B. A. Weinberg, and D. B. Mustard (2002): "Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997," *Review of Economics and Statistics*, 84(1), 45–61.

Grogger, J. (1998): "Market Wages and Youth Crime," *Journal of Labor Economics*, 16(4), pp. 756–791.

Hjalmarsson, R., and M. J. Lindquist (2010a): "Like Godfather, Like Son: Exploring the Intergenerational Nature of Crime," *Working Paper*.

——— (2010b): "The Origins of Intergenerational Associations in Crime: Lessons from Swedish Adoption Data," *working paper*.

Holmlund, H. (2007): "A Researcher's Guide to the Swedish Compulsory School Reform," *Swedish Institute for Social Research, Working Paper 9/2007*.

Holmlund, H., M. Lindhal, and E. Plug (forthcoming): "The Causal Effect of Parent's Schooling on Children's Schooling: A Comparison of Estimation Methods," *Journal of Economic Literature*.

Lochner, L. (2004): "Education, Work, and Crime: A Human Capital Approach," *International Economic Review*, 45(3), pp. 811–843.

Lochner, L., and E. Moretti (2004): "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *The American Economic Review*, 94(1), pp. 155–189.

Machin, S., O. Marie, and S. Vujić (2010): "The Crime Reducing Effects of Education," *CEP Discussion Paper No 979*.

Machin, S., and C. Meghir (2004): "Crime and Economic Incentives," *The Journal of Human Resources*, 39(4), pp. 958–979.

Marklund, S. (1980): "Från reform till reform: Skolsverige 1950-1975, Del 1, 1950 års reformbeslut," *Skolöverstylrelsen och Liber Utbildningsförlaget.*

——— (1981): "Från reform till reform: Skolsverige 1950-1975, Del 2, Försöksverksamheten," *Stockholm: Utbildningsförlaget.*

Meghir, C., and M. Palme (2003): "Ability, Parental Background and Education Policy: Empirical Evidence from a Social Experiment," *IFS working paper*, (WP03/05).

——— (2005): "Educational Reform, Ability, and Family Background," *The American Economic Review*, 95(1), pp. 414–424.

Moulton, B. R. (1990): "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *The Review of Economics and Statistics*, 72(2), pp. 334–338.

National School Board, R. (1959): "Försöksverksamhet med nioårig skolplikt," Stockholm, Report 52.

Sabates, R., and L. Feinstein (2008a): "Educational Attainment and Juvenile Crime," *British Journal of Criminology.*

——— (2008b): "Effects of Government Initiatives on Youth Crime," *Oxford Economic Papers*, 60(3), pp. 462–483.

Schweinhart, L., J. J. Montie, Z.-p. Xiang, W. S. Barnett, C. R. Belfield, and M. Nores (2005): *Lifetime effects: The High/Scope Perry Preschool Study Through Age 40.* Ypsilanti, MI: High/Scope Press.

Williams, J., and R. C. Sickles (2002): "An Analysis of the Crime as Work

Model: Evidence from the 1958 Philadelphia Birth Cohort Study," *The Journal of Human Resources*, 37(3), pp. 479–509.

8 APPENDIX

8.1 Reform Appendix

Table 12: Quantitative development of the comprehensive school experiment 1949 to 1962.

Year	Municipalities		Number of classes	Number of students
	Cumulative Number	Percentage share		
1949/50	14	1.3	172	2 483
1950/51	20	1.9	379	7 529
1951/52	25	2.4	682	14 635
1952/53	30	2.9	1 009	22 725
1953/54	37	3.5	1 525	35 784
1954/55	46	4.4	2 516	61 498
1955/56	59	5.6	3 394	84 941
1956/57	71	6.7	4 393	109 694
1957/58	96	9.1	5 702	143 370
1958/59	142	13.5	8 036	196 343
1959/60	217	20.6	11 191	266 042
1960/61	295	28.0	14 283	333 094
1961/62	415	39.4	18 665	436 595

Note: The 1952 division of municipalities (total: 1 052). Source: Marklund

8.2 Data appendix

Table 13: The Sample

	Number observations		
	All	Male	Female
Parent generation:			
Cohorts 1945-1955	1,340,857	685,056	655,801
Father's education available	881,742	452,433	429,309
Of which low educated father	560,273	287,396	272,877
Percent	63.54	63.52	63.56
Children generation:			
All children of parent generation	1,621,758	833,564	788,194
Paternal grandfather's education available	802,451	412,619	389,832
Of which with low educated grandfather	511,980	263,319	248,661
Percent	63.80	63.82	63.79
Maternal grandfather's education available	836,632	430,357	406,275
Of which with low educated grandfather	538,228	276,779	261,449
Percent	64.33	64.31	64.35

Notes: We only present the number of observations that are available on father's and grandfather's education level, because we will condition on father's or grandfather's education level in the analysis. We only have information on the highest level of education for those individuals that are not older than 60 years in the year of the 1970 census. We report the number of individuals in each sample, the number of individuals for whome we have information on the highest level of education on their fathers or grandfathers and the share of those for which we have this information with the lowest education level. For the children generation with low educated grandfathers on their father's side of the family, we consider those children whose father was born between 1945 and 1955. For the children generation with low educated grandfathers on the mother's side of the family we consider those whose mother was born between 1945 and 1955.

Table 14: Number of all convictions in Sweden between 1981-2008

	All	Male	Female
Number of convicted persons	1,249,569	966,790	282,779
Number of persons convicted for prison	366,639	344,919	21,720
Number of convictions in total	3,014,811	2,534,337	480,474
Number of prison sentences in total	1,204,711	1,115,428	89,283
Convictions by age group			
age 15 -24	1,128,125	950,413	177,712
age 25-34	710,177	605,445	104,732
age 35-44	577,693	483,821	93,872
age 45-54	355,396	296,971	58,425
age 55-64	161,367	133,788	27,579
age 65-80	76,296	59,138	17,158
age $>$ 80	5,757	4,761	996

78% to 85% of the convictions are males.

Table 15: Data on all suspected crimes in Sweden between 1991-2009. Sons of men or women born 1945-1955.

	Number of persons suspected for a crime 1991-2009		
	All crimes	Excluding traffic	Excluding some traffic
Sample: Sons of men born 45-55			
	129,683	117,279	124,487
Percent of sample	20.95	18.94	20.11
Sample: Sons of men born 45-55 with low educated father			
	54,542	48,888	52,222
Percent of sample	20.71	18.57	19.83
Sample: Sons of women born 45-55			
	133,953	120,748	129,217
Percent of sample	20.50	18.48	19.78
Sample: Sons of women born 45-55 with low educated father			
	55,210	49,294	129,217
Percent of sample	19.95	17.81	19.78

Notes: The category excluding all traffic crimes excludes the crime categories listed in table 17. The category excluding some traffic crimes excludes the categories "Driving without a license", "Allowed driving without license" and "Override provision".

Figure 2: Life cycle convicted crimes

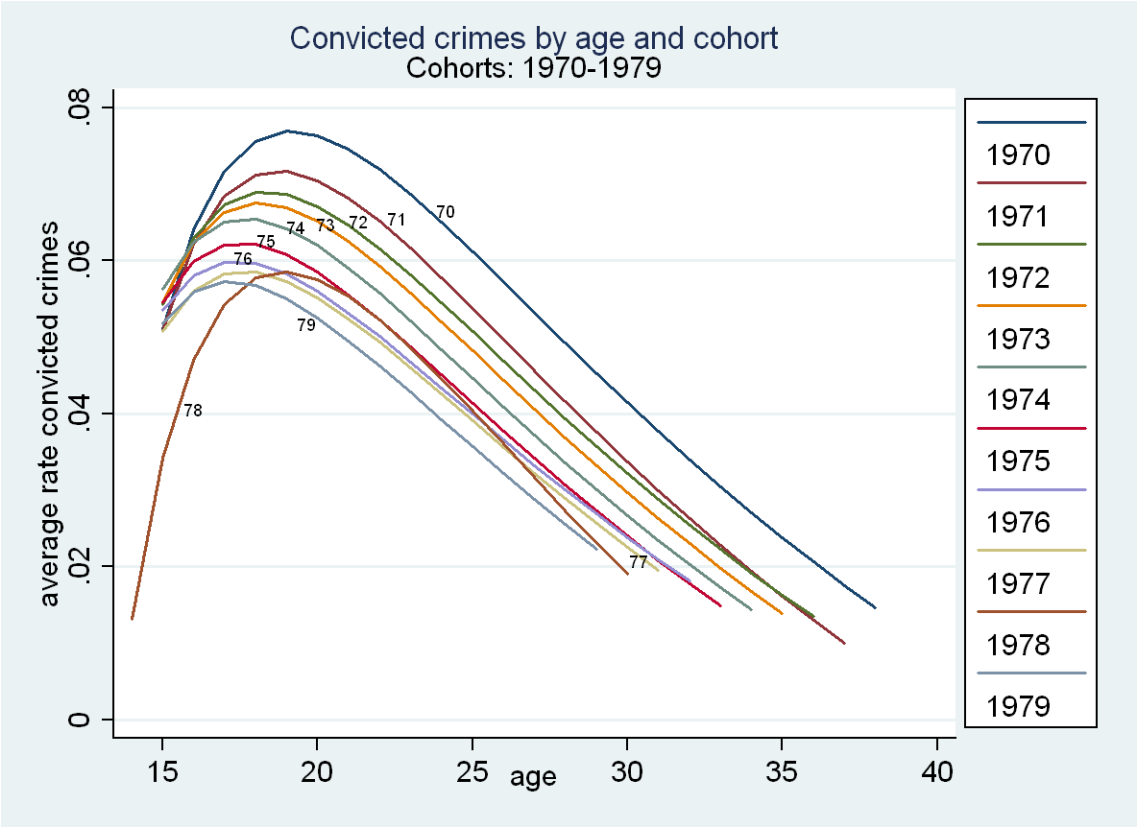


Figure 3: Life cycle convicted crimes

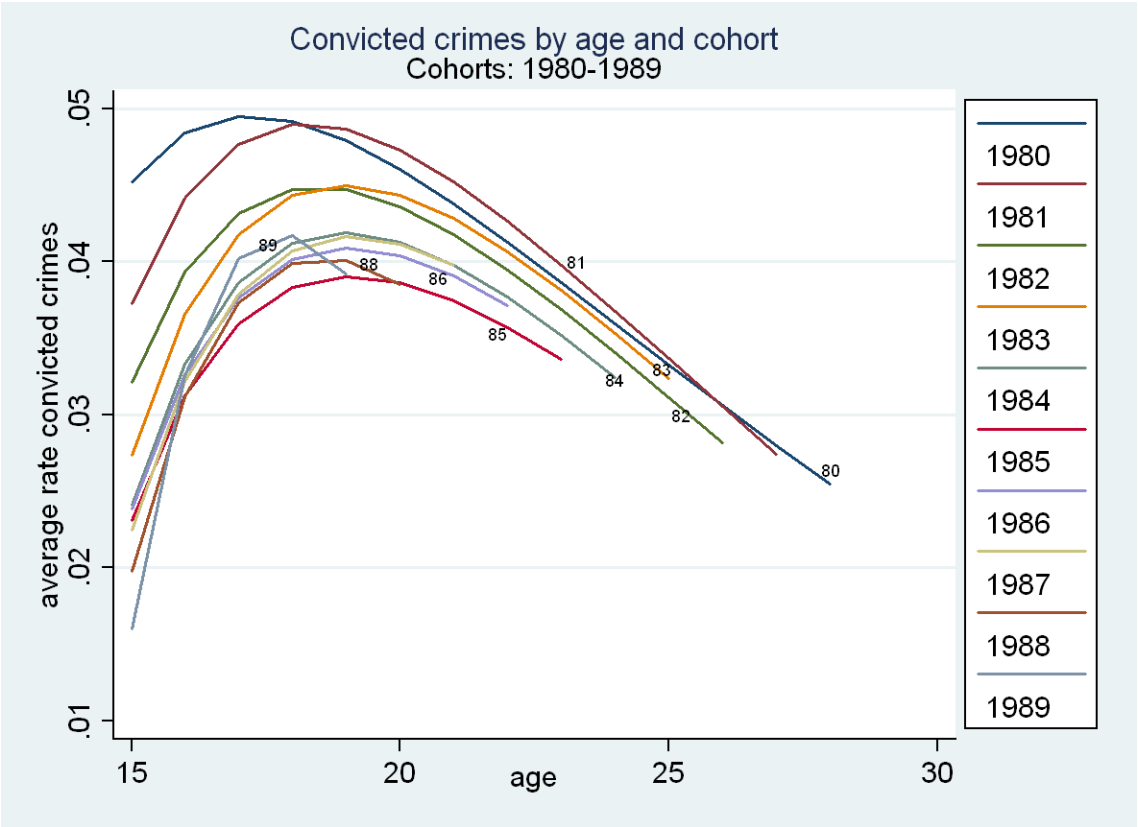


Table 16: Persons found guilty of criminal offences, by principal offence

	Number of convictions, 2009	Share, %
Crimes against penal code	**59,542**	**42.1**
Of which		
Crimes against life and health	*9,744*	*6.9*
Of which		
Murder and man-slaughter	150	0.1
Assault, gross assault	9,268	6.5
Sexual offences	*1,090*	*0.8*
Of which:		
Rape	256	0.2
Theft, robbery, other offences of stealing	*29,393*	*20.8*
Of which:		
Theft, gross theft	9,233	6.5
Petty theft	17,953	12.7
Robbery, gross robbery	1,049	0.7
Vehicle theft	824	0.6
Fraud and other dishonesty	*3,175*	*2.2*
Crimes inflicting damage	*3,316*	*2.3*
Violent threat to public servant	*2,544*	*1.8*
Other	*10,280*	*7.3*
Crimes to other penal legislation	**82,035**	**57.9**
Crimes against the Road traffic offences act	*47,020*	*33.2*
Of which		
Drunken driving, gross drunken driving	13,253	9.4
Crimes against the Narcotics drugs act	*18,525*	*13.1*
Crimes against the Act on smuggling	*2,076*	*1.5*
Other	*14,414*	*10.2*
All crimes	**141,577**	**100**

Persons found guilty of criminal offences, by principal offence, 2009. Source: Kriminalstatistik, Rättsstatistisk årsbok, Statistisk årsbok, Statistiska Meddelanden (R 11 SM).

Table 17: Traffic crimes persons were suspected for, January 1991 - June 2009.

Total number of crimes	**4,073,985**	
Total number of traffic crimes	**687,522 (16.876%)**	
Description of traffic violation	**Number of cases**	**Percentage of traffic crimes**
Driving or aggravated driving without a license	423,809	61.64
Drink-driving or aggravated drink-driving under the influence of alcohol alone or under the influence of both alcohol and drugs	167,958	24.43
Driving or aggravated driving under the influence of drugs alone	49,828	7.25
Hit and run	19,346	2.81
Allowed driving without a license	9,941	1.45
Gross negligence in traffic	9,018	1.31
Cause of danger for another in connection with traffic or in traffic	3,896	0.57
Causing bodily injury or disease related to traffic accident	1,393	0.2
Involuntary manslaughter in connection with accident	1,201	0.17
Override provision	414	0.06
Other traffic offense, imprisonment in range of penalties	321	0.05
Maritime Act: Drunkenness	314	0.05
Other crimes against maritime law, imprisonment in the range of penalties	83	0.01

Notes: Criminal offences related to traffic violations persons were suspected for. Register data on all crimes individuals have been suspected for between January 1991 and June 2009 in Sweden.

Table 18: Linear probability estimates of the association of own education levels and criminal behavior. Men born between 1945-1955.

Dependent variables	Probability conviction $\bar{p} = 0.2531$	Probability prison $\bar{p} = 0.0538$
Education Levels		
Vocational	-0.161	0.031
	(1.190)	(0.273)
Upper secondary	-7.471***	-3.928***
	(1.028)	(0.287)
Upper secondary + $\geq$ 1 year	-10.549***	-5.113***
	(0.913)	(0.288)
College/University	-13.782***	-5.929***
	(0.923)	(0.395)
PhD	-19.759***	-7.183***
	(0.713)	(0.545)
Observations	684,625	684,625

Notes: Results are scaled by 100. Level of observation are men born between 1945 and 1955 in Sweden. Robust standard errors in parentheses, clustered by birth municipality. Significance levels *** p<0.01, ** p<0.05, * p<0.1. All regressions include a full set of birth cohort dummies and birth municipality dummies.

Table 19: Linear probability model estimates of the association of parental education levels and own criminal behavior. Sons of parents born 1945-1955.

Dependent variables	Probability conviction $\bar{p} = 0.2645$	Probability prison $\bar{p} = 0.0343$
Panel A: Education levels father		
Vocational	-2.075***	-0.694***
	(0.214)	(0.088)
Upper secondary	-8.083***	-2.342***
	(0.339)	(0.142)
Upper secondary + $\geq$ 1 year	-9.719***	-2.457***
	(0.388)	(0.135)
College/University	-12.535***	-2.900***
	(0.463)	(0.214)
PhD	-13.829***	-3.029***
	(0.551)	(0.303)
Observations	754,121	754,121
Panel B: Education levels mother		
Vocational	-4.356***	-1.556***
	(0.291)	(0.102)
Upper secondary	-8.119***	-2.648***
	(0.473)	(0.181)
Upper secondary + $\geq$ 1 year	-10.015***	-2.854***
	(0.381)	(0.173)
College/University	-12.324***	-3.242***
	(0.508)	(0.230)
PhD	-14.553***	-3.443***
	(1.059)	(0.390)
Observations	754,121	754,121

Notes: Level indicates a dummy for each education level. The omitted education level is the lowest education level combined levels 1 and 2. Level of observation are sons of parents born between 1945 and 1955 in Sweden. Robust standard errors in parentheses, clustered by birth municipality. Significance levels *** p<0.01, ** p<0.05, * p<0.1. All regressions include a full set of birth cohort dummies and birth municipality dummies.