ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Landersø, Rasmus; Nielsen, Helena Skyt; Simonsen, Marianne

Working Paper

School starting age and crime

IZA Discussion Papers, No. 7228

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Landersø, Rasmus; Nielsen, Helena Skyt; Simonsen, Marianne (2013) : School starting age and crime, IZA Discussion Papers, No. 7228, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/71664

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DISCUSSION PAPER SERIES

IZA DP No. 7228

School Starting Age and Crime

Rasmus Landersø
Helena Skyt Nielsen
Marianne Simonsen

February 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

School Starting Age and Crime

Rasmus Landersø
Rockwool Foundation Research Unit
and Aarhus University

Helena Skyt Nielsen
Aarhus University
and IZA

Marianne Simonsen
Aarhus University
and IZA

Discussion Paper No. 7228
February 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 7228
February 2013

ABSTRACT

School Starting Age and Crime*

This paper investigates the effects of school starting age on crime while relying on variation in school starting age induced by administrative rules; we exploit that Danish children typically start first grade in the calendar year they turn seven, which gives rise to a discontinuity in children's school starting age. Analyses are carried out using register-based Danish data. We find that higher age at school start lowers the propensity to commit crime, but that this reduction is caused by incapacitation while human capital accumulation is unaffected. Importantly, we also find that the individuals who benefit most from being old-for-grade are those with high latent abilities whereas those with low latent ability seem to be unaffected by being old-for-grade in school.

JEL Classification: I21, K42

Keywords: old-for-grade, school start, criminal charges, violence, property crime

Corresponding author:

Helena Skyt Nielsen
Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
8210 Aarhus V
Denmark
E-mail: hnielsen@econ.au.dk

* Comments from seminar participants at BeNA Humboldt University, ISER, University of Essex, University of Sussex, KORA, the workshop on "Economics of Successful Children" at Aarhus University as well as participants at the DGPE workshop are appreciated. The usual disclaimer applies.

I. Introduction

This paper investigates long-term effects of school starting age on crime. A large literature has investigated effects of school starting age on test scores and has convincingly shown that starting school later leads to improved test scores. Black, Devereux and Salvanes (2011) and Crawford, Dearden and Meghir (2006) refine this type of analysis and show that the result is completely driven by an age-at-test effect: children who start school later are simply older when they perform tests and this leads to better performance. But much less is known about behavioral consequences of school starting age. Black et al. (2011) show that higher school starting age leads to improved mental health (for boys) and a lower risk of teenage pregnancies (for girls), while there is conflicting evidence regarding the risk of receiving ADHD diagnoses (Dalsgaard, Humlum, Nielsen and Simonsen (2012); Elder (2010); Evans, Morrill and Parente (2010)). Our paper contributes to this literature by considering outcomes related to criminal behavior. In addition, we investigate heterogeneity in effects across different types of crimes, timing of crime throughout the week, across ages and across the ability distribution.

As in Black et al. we rely on exogenous variation in the school starting age generated by administrative rules. In particular, we exploit that Danish children typically start first grade in the calendar year they turn seven, which gives rise to a fuzzy regression discontinuity design. By comparing children born in December with children born in January we investigate the effects of starting first grade at the age of 6.6 compared to 7.6. The identifying assumption is, of course, that these two groups do not vary with regards to other characteristics.

Our analysis uses Danish register-based data for children born in the period from 1981-1993 with crucial information on exact birth dates, a range of crime outcomes, and a rich set of background characteristics.

We find that higher age at school start lowers the propensity to commit crime and extend our analysis with an estimation of marginal treatment effects. Importantly, we find that the individuals who benefit most from being old-for-grade are those with high latent abilities whereas those with low latent ability seem to be unaffected by starting later in school. Furthermore, we investigate potential mechanisms behind these results: incapacitation seems to play an important role. For boys, being old-for-grade reduces criminal charges until age 17, and the effect is driven by property crime and crime committed on weekdays. For girls, being old-for-grade postpones initiation of crime until after age 19, and the effect is driven by violent crime and crime committed on Fridays. We find that the relative age of classroom peers does not seem to be behind the reduction in crime. Finally, we simulate different policy changes and estimate policy relevant treatment effects of these.

The paper is structured as follows: Section II briefly reviews the Danish institutional set-up and discusses mechanism through which child behavior may be affected by school starting age and Section III describes the methodology. Section IV presents our data, Section V the results and finally Section VI concludes.

II. Institutional setting and mechanisms

A. Educational Institutions and School Starting Age

During the period relevant for this study and up until 2009, Danish law stipulated that education was compulsory from the calendar year of the child's 7th birthday and until completion of 9th grade.[1] This school system is fortunate for a study like this one because there is no automatic relationship between the school starting age and minimum required schooling as there would be in the US and

[1] See the Danish Law of public schools.

the UK systems, for instance. After 9th grade, education was voluntary and could follow an academic path (starting with high school) or a vocational path (starting with vocational school).

The year before entering first grade, children could enroll in a voluntary preschool class. The preschool class, compulsory schooling from 1st to 9th grade as well as post-compulsory schooling were free of charge. Furthermore, most children below the age of six were inscribed into some form of public child care, which was heavily subsidized.[2]

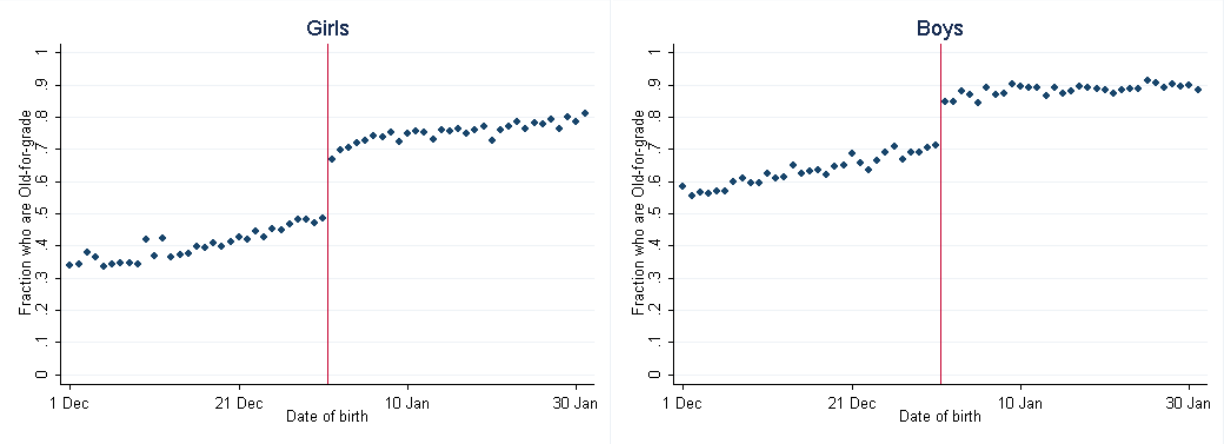
Because the formal age at school start is defined by the year of birth, each January 1st provides a cut-off point at which children born on each side are subject to a one year difference in timing of school start, even though they are born very close in time. Some parents of children born close to this cut-off date do choose to manipulate their children's' actual school starting age: children born at the end of the year are more likely to postpone school start one year, whereas children born early in the year are more likely to start school one year earlier than the law stipulates.[3] In consequence, some children born in December will start school one year later than they are supposed to - approximately at age 7.6 years - whereas the remainder of the children born in December will start when their age is around 6.6 years. Likewise, some children born in January will start school at age 6.6, which is one year earlier than the law stipulates, while the remainder will start school at age 7.6. Thus, school starting age for children born around the cut-off date is effectively reduced to a binary outcome: either children start at age 6.6 or they start at age 7.6. Figure 1 shows the fraction of children who are old-for-grade (age≈7.6) by date of birth for each gender.

[2] A minimum of 75 % of the expenses is covered by the local authorities, c.f. the Danish Law of day care.

[3] A recent white-paper on school start concluded that "many parents worry whether their children are ready to start school, and these concerns are supported by the preschool staff", cf. God Skolestart (2006).

FIGURE 1

FRACTION WHO ARE OLD-FOR-GRADE BY DATE OF BIRTH



Note: Being old-for-grade implies that the child starts school at age 7.6 instead of at age 6.6.

We see that there is a smooth upward trend in the fraction of girls and boys who are old-for-grade in December followed by a large discontinuity around New Year. The figure also shows that boys are more likely than girls to be old-for-grade.

School starting age may affect later behavior through skill formation. Cunha and Heckman (2008) show that cognitive and (especially) non-cognitive skills at pre-school ages are key determinants of later skill acquisition, behavior, and adult outcomes. If different school starting ages are associated with different levels of skills ("school readiness"), then these differences may be amplified and affect both other outcomes such as the tendency to engage in criminal activities.

Another mechanism by which school starting age may affect criminal activity is the incapacitation effect (see Lochner, 2011). When youth are in school, they simply have less time to commit crime. Previous studies confirm that this mechanism is important although complex. For urban areas, Jacob and Lefgren (2003) and Luallen (2006) find that increasing the number of school days reduce arrests for property crimes, while it increases arrests for violent crimes. While the effect on property crime is thought to be due to incapacitation, the effect on violent crime is thought to be a network

effect (spending more time with criminal others). Anderson (2012), which is closer to our approach, finds that increasing the compulsory school *leaving* age reduces arrests for both violent and property crime. Anderson (2012) estimates that a minimum school dropout age of 17 (relative to 16 or lower) decreases arrest rates for 17 year-olds by 8 %. This effect may be due to incapacitation, but strictly speaking, it is observationally equivalent to a situation where crime reports to the police differ according to whether youth are in school or not. One could imagine that criminal events taking place in school are treated differently than criminal events taking place outside school, which would lead to similar findings. This issue may be more relevant for violent crime than for property crime or traffic incidents.

A third mechanism through which school starting age may affect crime is by changing the individual's placement in the age hierarchy. Increasing school starting age by one year will most likely move the individual from being one of the youngest to being one of the oldest children in his/her classroom (e.g., Gaviria and Raphael (2001) and Sacerdote (2001)). However, the implications of such a change are ambiguous. On the one hand, having older peers who are more likely to engage in risky behavior may spark risky behavior at an earlier stage. On the other hand, having older peers might also increase skill acquisition and maturity, thus lowering risky behavior and improving educational outcomes. Fredriksson and Öckert (2005) and Black, Devereux and Salvanes (2013) find no substantial impact of the age composition of peers on educational and labor market outcomes or on teenage pregnancies among girls, and thus rule out that the relative age composition in class explains the impact of school starting age.

We formally investigate the importance of the incapacitation effects and the relative age of peers in our empirical analysis below.

B. Institutions guarding juvenile crime

Below we describe the institutions that may be relevant for understanding the potential impact of schooling and school starting age, in particular, on criminal activity of teenagers.

In Denmark, the age of criminal responsibility is 15, which is high in an international comparison; England has an age of criminal responsibility of 10, while only few US states have a limit and in those cases the limit is 6-12 years.[4] Until age 15, Danish children cannot be arrested, brought to court or imprisoned, although they may be withheld up to 6 hours by the police in which case a social worker must be present during interrogation. At ages 15-17, youth are considered fully responsible for their criminal acts, and may be imprisoned, though this should be separate from adult prisoners.[5] Thus, the focus is on prevention and rehabilitation rather than prosecution and punishment.

All local authorities have an interdisciplinary framework for prevention of juvenile crime involving the schools, the social services and the police (denoted SSP). This is a network of relevant players who collaborate to understand and prevent juvenile crime in the local area. They are concerned with general, specific as well as individual-oriented policies and interventions.

Reported victimization rates in Denmark are falling like in the rest of the OECD. However, overall crime rates in Denmark are somewhat higher than Norway, Sweden and the OECD average (19 v. 16 %) while they are almost at par with the US (18 %) and the UK (21 %), see OECD (2009). Therefore, we have no particular reason to expect that the effects of school starting age on crime should be substantively different in Denmark compared to other countries.

[4] http://www.unicef.org/pon97/p56a.htm.

[5] See the Danish Service Act.

III. Methodology

Our goal is to estimate the effect of school starting age (*SSA*) on crime outcomes (*Crime*):

$$Crime_i = X_i^{'}\beta_1 + \beta_2 SSA_i + \varepsilon_i,$$

where X is a conditioning set that includes a constant and child and parental characteristics predictive of *SSA* and *Crime*: APGAR score, birth weight, gestation length for children, mothers age at the birth of first child, both parents' education and labor market participation, a flexible function of distance in days to the cut-of and a constant. β_2 is our parameter of interest.

To circumvent the problem that *SSA* is not random and likely related to *Crime* through ε_i, we employ a strategy similar to Black et al. (2011), Evans, Morrill and Parente (2010), and Elder (2010). In particular, we exploit that school starting rules imply that children born just prior to January 1st are on average younger when they enroll in school than children born immediately after January 1st. We can therefore instrument *SSA* with a dummy for being born immediately after January 1st. As argued by Black et al. (2011), Evans et al. (2010) and Elder (2010), such cut-off dates constitute valid instruments in the sense of being uncorrelated with unobserved characteristics of child outcomes. In practice, we consider a short bandwidth with children born ± 30 days around January 1st. Results are robust to reducing or extending this bandwidth and to 'donut hole' strategies where observations close to the cut-off point are left out of the analysis.

Our endogenous variable of interest is essentially a dummy for starting school at age 7.6 relative to 6.6. Denote this new variable old-for-grade (*OG*) and suppress subscript *i*. Using a standard selection equation we may describe *OG* as:

$$OG = 1[\mu(X, z) - u > 0] \quad (1)$$

Where z is an indicator for being born before the cut-off date at January 1st: $z = 1[birth = December]$. Hence, we get the following first stage equation:

$$OG = X\gamma_1 + \gamma_2 1[birth = December] + u$$

Because z constitutes an exogenous instrument for being old-for-grade and plausibly assuming that no one starts school at earlier ages if born in January instead of December, we can estimate the Local Average Treatment Effect of *OG* as:

$$\beta_2^{LATE} = E\{\beta_2 | X\gamma_1 < u < X\gamma_1 + \gamma_2\}$$

This will capture the average effect of being old-for-grade for the group of children who would be inclined to change their school starting age, had their month of birth been different. This group may include children with very different sets of pre-school abilities and other pre-school characteristics. Thus β_2^{LATE} may comprise substantial differences across children at different margins. To uncover any differences across different levels of abilities we estimate the Marginal Treatment Effects (MTE) of being old-for-grade; see Björklund and Moffitt (1987) and Heckman and Vytlacil (2005). We define the MTE as:

$$\beta_2^{MTE} = E\{\beta_2 | v = v^*, X = X^*\}$$

where $v = F_u[u]$.[6] The MTE is constructed by using the instruments locally to estimate the effect of being old-for-grade for those who are on the margin of treatment at different points of v. Following the selection equation (equation 1), the margin of treatment is the point where treatment probability

[6] In terms of Marginal Treatment Effects the estimate of the 2SLS model will equal:

$$\beta_2^{LATE} = \frac{1}{U_D - U_{D^*}} \int_{U_D}^{U_{D^*}} \beta_2^{MTE} dU$$

where U_{D^*} is the random variable for those who are affected by the instrument.

by observed characteristics $\mu(X, z)$ must equal latent characteristics u (v). So an individual with large observed treatment probability and located on this margin (i.e. an individual with low levels of observed abilities) must have an equally large u (v) to offset the disadvantageous observable characteristics and make $\mu(X, z) = u$. Oppositely, individuals on the margin of treatment with low observed treatment probabilities must have small values of u . $v = F_u[u]$ can therefore be interpreted as a scale measuring increasing levels of latent abilities. Hence, the MTEs allow us to distinguish between heterogeneous treatment effects across latent abilities v.[7]

IV. Data

We use Danish register-based data for children born in the period from mid 1981- mid 1993 with crucial information on exact birth dates, charges[8] for property crime, violence and other types of crime (in particular traffic incidents), together with the specific dates of crime, and the usual set of background characteristics.

Using these registers we combine information on the children's birth weight, gestational length, APGAR score, demographic variables, educational variables, and crime by the unique individual identification number. We also have information on parents' identification number. Using these we can link parents' characteristics to their children and also identify whether the individuals in question have had any children themselves. For parental background characteristics we use education and labor market outcomes as measured one year prior to the child's birth. Importantly,

[7] We estimate the Marginal treatment effects using a parametric setup as outlined in Heckman, Urzua, and Vytlacil (2006). Our results are robust to alternative specifications as e.g., a semi parametric estimation.

[8] We would have appreciated being able to look at convictions. However, to obtain an adequate measure of convictions one needs to wait three years after the charge which would reduce our data material substantially.

we center all covariates and outcome variables around the cut-off dates instead of by calendar year. Hence, we compare information on children born in January year t to the information on children born in December year t-1 instead of comparing information on children born in January year t to the information on children born in December year t.[9]

Unfortunately, we do not have information on the specific timing of school starting age for the cohorts in question. Instead we use age in 8th grade as an approximation. We do observe children's exact ages at all grade levels from 2007 and onwards and we use this data to check that the approximation of school starting age by age in 8th grade works very well (see Table A1). The vast majority of children who have not completed elementary school by age 15 do so because they are old-for-grade already in preschool class, while very few children are delayed from the first grade and onwards.

Table A2 in Appendix shows mean background characteristics for children born 30 days before and after January 1st for boys and girls separately. We see that some differences are significantly different but all are very small in size and often vary by gender.[10] We include these variables as covariates and experiment with further restrictions on the birth date intervals, and our results are robust to these restrictions.

Our main outcomes measure age-specific criminal charges until the 18th birthday, both aggregate crime charges and specific types of crime (property crime and violent crime). To investigate the age profile of the impact of school starting age more closely, we also consider criminal charges two

[9] For children born in December 1981 or January 1982 we use parental characteristics measured in 1980, while we for children born in December 1982 or January 1983 use parental characteristics measured in 1981 etc.

[10] The difference in birth weight is for example 16 grams, which corresponds to 0.03 point difference in IQ according to Black, Devereux and Salvanes (2007).

years after the completion of 9th grade as well as accumulated age-specific criminal charges until ages 16-21. In our mechanism analysis, we also consider completed months of schooling by age 27.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR MAIN OUTCOME VARIABLES

Variable	Girls		Boys	
	December	January	December	January
Criminal charges by 18th birthday (0/1)	**0.057**	**0.051**	**0.196**	**0.181**
	0.232	0.219	0.397	0.385
- For property crime	0.043	0.040	**0.089**	**0.080**
	0.202	0.196	0.285	0.271
- For violent crime	**0.006**	**0.004**	0.019	0.017
	0.077	0.061	0.135	0.129
- Other crime	0.008	0.007	0.088	0.085
	0.091	0.083	0.284	0.279
Criminal charges 2yrs after graduation from compulsory school (0/1)	0.046	0.044	0.151	0.153
	0.210	0.204	0.358	0.360
# Observations	48,546	48,546	50,383	50,383

Note: Bold indicates that the mean for children born in December and January are significantly different at a 5%-level.

Table 1 shows mean crime outcomes by birth-month and gender. We see that around 5% of girls and almost 20% of boys have been charged with a crime prior to their 18th birthday. For girls, the dominant charge is for property crime. For boys around 8-9% have been charged with a property crime, around 2% have been charged with a violent crime, and around 8-9% have been charged with other types of crime, mainly traffic violations. Also, there are some differences in the crime rates across birth month. Those born in December are more likely to have been charged with a crime compared to those born in January.

V. Results

A. Timing of birth within the calendar year and school starting age

Table 2 presents our first stage results, using a dummy for birth in December as instrument for old-for-grade (*OG*) together with year specific December dummies to allow for trends in the fraction of children who are old-for-grade over the period. The table shows the first stage results estimated both with and without controls.

In line with Figure 1, we see that the instrument strongly predicts whether children start school at age 7.6 or 6.6: children born in January are significantly more likely to be relatively old when they start school compared to children born in December and the effect is large. This is despite the tendency for some children born in December to delay enrolment and start at age 7.6 instead. The cut-off identifiers are highly significant (t-value around 25 for girls and around 17 for boys) and the associated F-statistics pass the Staiger-Stock rule-of-thumb.[11]

B. Crime results

Table 3 shows our main estimation results. We find that being old-for-grade leads to significant reduction in the propensity to commit crime before age 18, regardless of gender. Effects are large relative to the mean. The share of girls with any criminal charges at age 18, for example, is 0.054 among children born ± 30 days around January first. The effect of starting school at age 7.6, in comparison, is 1.5 percentage points reduction, or just below 30 % of the mean. For boys, the effect of school starting at age 7.6 on criminal charges at age 18 is a 4 percentage point reduction, which should be seen relative to a share of boys with criminal charges of 0.19. We also find that being old-

[11] With one endogenous variable and 12 instruments, F should be greater than 21.

for-grade predominantly lower girls' propensity to commit violent crime, whereas for boys it lowers the propensity to commit property crime.

TABLE 2

FIRST STAGE ESTIMATION RESULTS

CHILDREN BORN IN DECEMBER AND JANUARY

Variables	Girls		Boys	
	2SLS	2SLS	2SLS	2SLS
Born in Dec	-0.34 ***	-0.34 ***	-0.21 ***	-0.21 ***
	-24.99	-25.23	-17.46	-17.59
Born in Dec 1981	-0.03	-0.03	-0.15 ***	-0.15 ***
	-1.67	-1.61	-8.47	-8.54
Born in Dec 1982	-0.08 ***	-0.08 ***	-0.15 ***	-0.15 ***
	-3.72	-3.83	-8.47	-8.49
Born in Dec 1983	-0.06 **	-0.06 **	-0.09 ***	-0.09 ***
	-2.85	-3.04	-5.02	-5.02
Born in Dec 1984	-0.04 *	-0.05 *	-0.08 ***	-0.08 ***
	-2.10	-2.36	-4.67	-4.63
Born in Dec 1985	-0.04 *	-0.05 *	-0.08 ***	-0.08 ***
	-2.07	-2.22	-4.36	-4.55
Born in Dec 1986	-0.03	-0.03	-0.06 ***	-0.06 ***
	-1.35	-1.65	-3.44	-3.54
Born in Dec 1987	-0.02	-0.03	-0.05 *	-0.05 **
	-1.19	-1.25	-2.57	-2.60
Born in Dec 1988	0.01	0.01	-0.02	-0.03
	0.49	0.39	-1.36	-1.46
Born in Dec 1989	-0.03	-0.03	-0.04 *	-0.03
	-1.60	-1.69	-2.06	-1.93
Born in Dec 1990	-0.01	-0.01	0.01	0.01
	-0.43	-0.51	0.67	0.58
Born in Dec 1992	0.02	0.02	-0.01	-0.01
	1.03	0.85	-0.41	-0.40
Controls				
- Yearly cut-off fixed effects	X	X	X	X
- Covariates		X		X
F-value	368.73	234.46	280.35	170.60
# Observations	48,546	48,546	50,383	50,383

Note: 1991 is the reference year.* indicates $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

TABLE 3

ESTIMATION RESULTS

Variable	Girls			Boys		
	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Criminal charges by 18th birthday (0/1)	0.001	-0.016 **	-0.015 **	0.000	-0.051 ***	-0.036 **
	0.49	-2.9	-2.78	0.08	-4.00	-2.86
- For property crime	0.001	-0.008	-0.007	-0.001	-0.031 ***	-0.022 *
	0.63	-1.61	-1.46	-0.30	-3.39	-2.45
- For violent crime	0.000	-0.005 **	-0.005 **	0.000	-0.008	-0.006
	0.36	-3.11	-3.11	-0.23	-1.95	-1.45
- Other crime	0.000	-0.003	-0.003	0.002	-0.012	-0.007
	-0.44	-1.37	-1.33	0.52	-1.30	-0.82
Criminal charges 2 years after 9th grade (0/1)	0.01 ***	-0.006	-0.005	0.05 ***	0.01	0.02 *
	5.39	-1.2	-1.05	13.30	0.85	2.09
Controls						
- Yearly cut-off fixed effects	X	X	X	X	X	X
- Covariates			X			X
# Observations	48,546	48,546	48,546	50,383	50,383	50,383

Note: * indicates $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

Just as Black et al. (2011) show in the case of test scores, it is extremely important to align children in terms of age in analyses of crime outcomes. If one ignores this and considers criminal charges for example two years after the completion of 9th grade, we wrongly conclude that school starting age does not affect crime outcomes – but this is just because the tendency to commit crime and therefore the risk of criminal charges increase with age in the relevant age range. However, the age-gradient is surprisingly small; it nullifies the effect for girls and barely switches the sign to borderline significantly positive for boys.

C. Marginal treatment effects

This section presents the estimation results and confidence intervals for the marginal treatment effects of higher school starting age on crime until age 18. In all figures, the x-axis shows the normalized term *v* (see section III) which is (from the researcher's perspective) unobserved or latent ability. As *v* approaches 1, the level of latent ability increases. I.e. individuals with high values of *v* must be subject to worse conditions than individuals with low values of *v*, in order to be at the margin of being treated. Of course, we can only estimate the MTE within the intervals of common support. Figure A1 and A2 show these areas for girls (.15-.90) and boys (.35-.95). The lack of common support for low values of latent ability for boys arises because the least able always delay school start. The weighted average of the different MTEs across the entire set of latent ability equals the 2SLS estimates from Table 3. All confidence intervals have been non-parametrically bootstrapped. Figures 1 and 2 show the MTEs of being old-for-grade on crime for girls and boys, respectively. Since *Crime* is a dummy variable, an estimate of e.g. -0.01 implies that increasing school starting age by one year results in an reduction of the probability of a criminal charge before the age of 18 by 1 percentage point.

We see from Figure 1 that the estimated marginal treatment effects of being old-for-grade on crime for girls are negative and significant for almost the entire interval of support. The MTEs appear to be relatively constant across different values of latent ability. From Figure 2 we see that the MTE of being old-for-grade on crime for boys is downwards sloping across latent ability. The estimated effects at the low intervals of latent ability are insignificant but decreases (and increases in significance) as latent ability increases. We also note that the average MTE is around -0.015 for girls and -0.04 for boys which correspond to the 2SLS from Table 3. However, Figure 2 also shows that the 2SLS estimate for boys conceal considerable heterogeneity in the effects of being old-for-grade across different levels of ability, as being old-for-grade reduces crime for the most able boys,

whereas the crime rates of the least able boys appear unaffected by school starting age. It makes sense that the boys with the highest latent abilities benefit the most from being in school because they are supposedly more comfortable in that environment than the boys with lower latent abilities.[12]

[12] Inspired by the strategy by Black et al. (2011), we have rerun our main results from Table 3 for the subgroups of boys enrolled in high school, vocational school, and those not yet enrolled at age 18. We find that the negative impact of school starting age on crime for boys is driven by the impact on boys enrolled in high school, while the effect for boys enrolled in vocational school is insignificant. This may suggest that high school is more protective of crime than the alternative environments, while other explanations may also be relevant.

FIGURE 1

MARGINAL TREATMENT EFFECTS, CRIME AGE 18, GIRLS

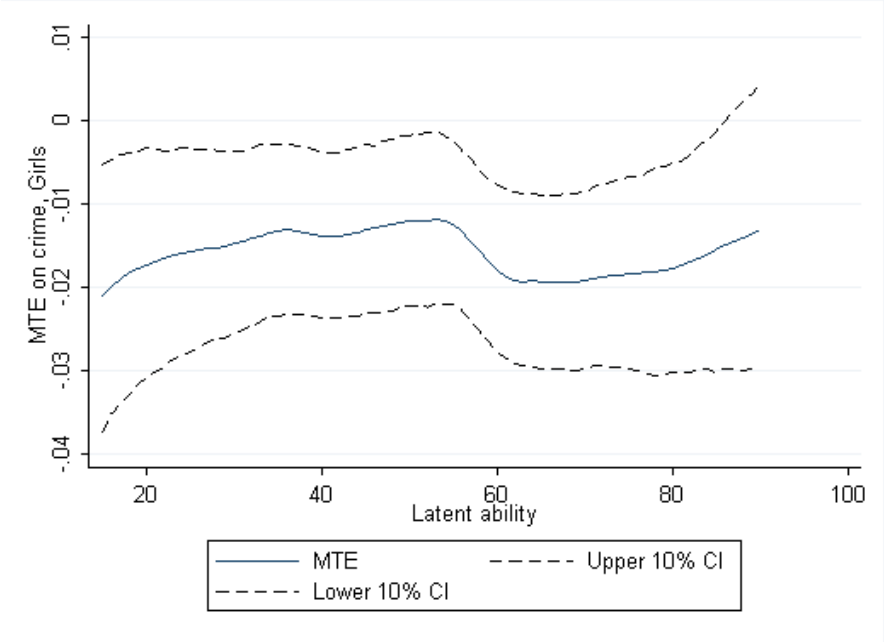
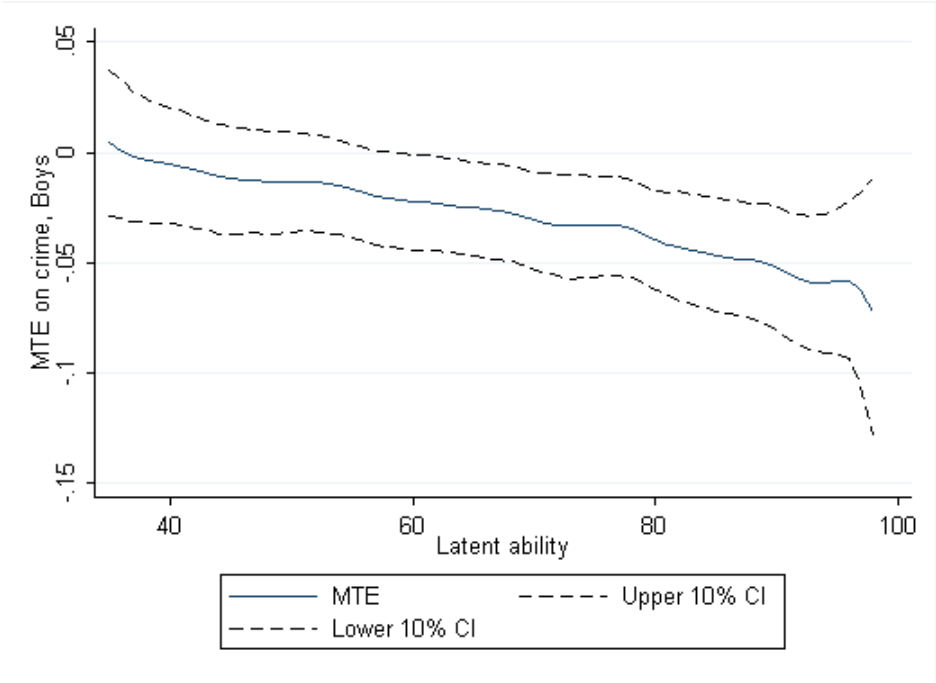


FIGURE 2

MARGINAL TREATMENT EFFECTS, CRIME AGE 18, BOYS



D. Potential Mechanisms

This section attempts to shed light on some of the different channels through which school starting age may affect crime outcomes. Specifically, we investigate the importance of incapacitation and the role played by the relative age of peers as in Black et al. (forthcoming).

Table 4 investigates the effects of school starting age using our main estimation strategy for criminal charges before ages 16-21. If incapacitation is important for our effects of *SSA* on *Crime* we should expect the effects of *SSA* to be larger at younger ages when children are more likely to actually be in school. For girls, this is exactly what we find. The effects on accumulated charges are significantly negative until age 19 when most have completed high school but then turn insignificant. Thus, school starting age influences the timing of charges but not the accumulated total. For boys, on the other hand, all point estimates are significantly negative, and the magnitude goes up from 0.022 before the 16th birthday to 0.034 before the 17th birthday, and then it stays almost constant until the age of 21. This indicates that boys who are old-for-grade have a slightly lower probability of being charged at age 15 and age 16, while they are still in school. However, since the effect on accumulated crimes stays constant after the 17 there is no difference after they leave comprehensive school. However, there might still be a permanent effect if early charges are kept in the police records.

Figures A5 and A6 illustrate the marginal treatment effects over ages 16-21. For girls the effects of school starting age on crime are negative over the whole range of ages and latent abilities, while for boys the effects are positive for the youngest ages for boys with low latent abilities.

In Table 5, we investigate how school starting age affects crime across the week.[13] For boys, the effect is driven by crime committed during the weekdays. For girls, the effect during the weekdays is only significant when Fridays are included. This is consistent with the earlier documented gender difference: the effect of school starting age on crime for boys is primarily driven by property crime while the effect for girls is driven by violent crime, which apparently, at least to some extent, takes place on Fridays.

We interpret the findings reported in Tables 4 and 5 as indicating that the effects for both gender are consistent with some type of incapacitation effect. For boys who are old-for-grade, we find that charges (primarily for property crime) are lower up until the 17th birthday while they are still in school, and this is due to fewer charges for crimes committed on weekdays. For girls who are old-for-grade, we see that observed charges (primarily for violence) are postponed, and in the long run the effect on accumulated charges is zero. These results comport with our findings on completed education at age 27 (not shown but available on request). The estimated effect on length of education is insignificant and practically zero for girls, while it is about two additional months for boys (significant at the 5 % level).

Table 6 analyzes the effect of the age of peers in line with Fredriksson and Öckert (2005) and Black et al. (forthcoming). Formally, we include the average age of peers in one's school in 8th grade as an additional control variable in our models of crime outcomes. To handle endogeneity of the average age of peers, we instrument with the predicted average age of peers *had everybody started on time*.[14] We see that the mean age of peers has no significant effect on crime outcomes and that the effect of own school starting age is completely unaffected by the inclusion of this extra control

[13] Table A3 shows the coefficient for each single day of the week. Most of these are insignificant,

[14] We impute average age of peers for observations with fewer than 10 other children enrolled at the school in grade 8.

variable. This is in line with the findings in the mentioned previous studies for Norway and Sweden.

TABLE 4

EFFECT OF SCHOOL STARTING AGE ON CRIME AT AGES 16-21

	Girls	Boys
Crime prior to 16th birthday	-0.010 *	-0.022 *
	-2.36	-2.49
# Observations	48,546	50,383
Crime prior to 17th birthday	-0.014 **	-0.034 **
	-2.77	-3.12
# Observations	48,546	50,383
Crime prior to 18th birthday	-0.015 **	-0.036 **
	-2.78	-2.86
# Observations	48,546	50,383
Crime prior to 19th birthday	-0.017 **	-0.039 **
	-2.62	-2.81
# Observations	43,668	45,368
Crime prior to 20th birthday	-0.010	-0.036 *
	-1.38	-2.33
# Observations	39,037	40,606
Crime prior to 21st birthday	-0.008	-0.037 *
	-1.04	-2.26
# Observations	34,559	36,012
Controls		
- Yearly cut-off fixed effects	X	X
- Covariates	X	X

Note: * indicates $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

TABLE 5

EFFECT OF SCHOOL STARTING AGE ON CRIME ACROSS THE WEEK

	Girls	Boys
Crime prior to 18th birthday		
Monday-Thursday	-0.006	-0.022 *
	-1.39	-2.23
Monday-Friday	-0.012 *	-0.030 **
	-2.33	-2.75
Saturday-Sunday	-0.003	-0.005
	-1.51	-0.73
# Observations	48,546	50,383

Note: * indicates p < 0.05, ** p < 0.01, and *** p < 0.001

TABLE 6

ABSOLUTE AND RELATIVE EFFECTS OF SCHOOL STARTING AGE ON CRIME

	Effect of old-for-grade		Effect of peer age	
	Girls	Boys	Girls	Boys
Crime (prior to 18th birthday)	-0.015 **	-0.037 **	0.003	0.067
	-2.78	-2.93	0.07	0.78
Property crime	-0.007	-0.023 *	0.018	0.070
	-1.49	-2.56	0.45	1.15
Violent crime	-0.005 **	-0.007	-0.003	0.037
	-3.10	-1.57	-0.24	1.26
Other crime	-0.003	-0.007	-0.012	-0.041
	-1.28	-0.75	-0.67	-0.65
Crime 2 years after grad from 9th	-0.005	0.026 *	-0.012	-0.032
	-1.04	2.26	-0.27	-0.42
Controls				
- Yearly cut-off fixed effects	X	X	X	X
- Covariates	X	X	X	X
# Observations	48,546	50,383	48,546	50,383

Note: * indicates p < 0.05, ** p < 0.01, and *** p < 0.001

E. Policy Relevant Treatment Effects

One question is how being old-for-grade vs. not being old-for-grade affects crime. In the previous sections we have assessed this, first by estimating the average effects for those who are affected by the cut-off date at New Years Eve and second by estimating the effects at different points of latent ability. In this section we will ask a different question, namely how different policy changes to school starting age would affect crime. Unlike the previous sections, the counterfactual is not "no-treatment" but instead the baseline policy. Also, the estimated effects depend on which group of individuals that are induced to change their school starting age by such a change.

We estimate Policy Relevant Treatment Effects (PRTE) on the outcomes of interest for those who would be inclined to change their treatment response as a result of three different policies for three different policy changes: 1) A 20 % increase in propensity to allow children to be old-for-grade, i.e. less enforcement of school starting laws, 2) moving the cut-off date four weeks forward to January 28th, and 3) moving the cut-off date four weeks back to December 4th.

In order to estimate the effects of these three policies we must identify to what extent individuals at different margins of treatment (i.e. at different points in the distribution of latent characteristics) change their treatment response. The reason is that an individual's treatment assignment is determined by whether the predicted probability of treatment through observable characteristics less one's latent characteristics is larger than zero (see equation 1). However, each individual may be at a different margin of treatment and may have a different predicted probability of treatment by observable characteristics for different policies. For instance if formal SSA was 9 years, then only the least able (low values of v) would be induced to be old-for-grade whereas if SSA was 3 years then only the most able (high values of v) would not be old-for-grade. Hence, for different policies, different sets of individuals (with different latent characteristics) will be old-for-grade. And as the

effects of being old-for-grade vary across v, one should re-weight the estimates so a large weight is assigned to the intervals in which many change their treatment response as a result of a policy change and vice versa. Heckman and Vytlacil (2005) show that PRTEs can be estimated as weighted averages of the MTEs:

$$\beta_2^{PRTE} = \int \beta_2^{MTE} \omega_i dU \quad (2)$$

Using $\omega_i = {F_P(U_D) - F_{P^*}(U_D)} \big/ {E(P^*) - E(P)}$ as weights where P is the predicted probability of being old-for-grade under the baseline policy and P^* is the predicted probability under the new policy.[15]

Table 7 shows the PRTEs of the three different policies on total crime at age 18 and at crime until graduation plus 2 years. Figure A3 in the appendix shows the PRTE weights for girls and figure A4 shows the corresponding for boys.[16]

[15] The integral from equation 2 depends on our support of v which ideally would cover the entire range from 0 to 1. Figures A1 and A2 in the appendix shows that we only have support in the intervals (0.02;0.99) for girls and (0.11;0.99) for boys. However, the margins of treatment and associated weights for the three new policies lie well within our interval of support from our baseline policy from which we get identification. In consequence, the lack of support does not disqualify the estimations though the PRTEs may be biased.

[16] We only show weights for intervals that correspond to the area of support, as the weights at the remaining intervals represent very few outliers and do affect the estimated PRTEs in practice.

TABLE 7

POLICY RELEVANT TREATMENT EFFECTS

	Policy 1		**Policy 2**		**Policy 3**	
	Girls	Boys	Girls	Boys	Girls	Boys
Criminal charges by 18th birthday (0/1)	-0.023 **	-0.068 **	0.025 ***	0.055 **	-0.024 ***	-0.076 **
	-3.08	-2.91	3.45	2.83	-3.32	-2.92
- For property crime	-0.010	-0.044 **	0.014	0.034 *	-0.013	-0.054 **
	-1.36	-2.65	1.90	2.43	-1.75	-2.75
- For violent crime	-0.007 **	-0.013	0.007 **	0.011	-0.008 **	-0.013
	-2.84	-1.48	2.67	1.58	-2.71	-1.39
- Other crime	-0.005	-0.007	0.004	0.006	-0.004	-0.006
	-1.54	-0.42	1.39	0.44	-1.38	-0.33
Controls						
- Yearly cut-off fixed effects	X	X	X	X	X	X
- Covariates	X	X	X	X	X	X
# Observations	48,546	50,383	48,546	50,383	48,546	50,383

Note: * indicates $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

From the table we see that Policy 1 and Policy 3, which will induce more students to be old-for-grade, will have a favorable *medium run* impact for the boys who go from being young-for-grade to being old-for-grade as a result of the policies. The reason is that these policies work through allowing more individuals with high latent abilities to start school later, and these individual are exactly the ones for whom we see strong favorable effects of starting school later (see Figure 2). Hence, the propensities to commit crime before the age of 18 will decrease as a consequence of these two policies. Similarly Policy 2, which will reduce SSA, will induce some boys who were old-for-grade to become young-for-grade. Again, the boys affected by the policy will be boys with high latent abilities, and therefore we see a strong unfavorable impact on the propensity to commit crime before the age of 18. These results reflect that the estimated MTEs are negative at the intervals where the absolute values of PRTE weights are largest. From Table 6 we also see that the impacts of the three policies are analogous for girls although effect sizes are smaller and reflect the homogenous pattern seen in Figure 2.

The estimates from Table 6 should be viewed as approximations of the effects of the three policy changes, as the estimates e.g. do not take general equilibrium effects to criminal opportunities into account. One such equilibrium effect would go through the relative age composition. For instance, the general equilibrium effect of changing the cut-off date by 4 weeks in either direction would be neutral if the effect was driven by the relative age composition, which however as mentioned earlier, is not the case. Another example of an equilibrium effect could be exacerbation of the effects if more students become violent or dampening of effects if the easy access to property is reduced when more students steal. Moreover, the PRTEs do not capture the overall costs and benefits of the various policies. A comprehensive cost-benefit analysis should on the one hand take the *long run* individual effects of increased educational attainment, lower crime rates but also later entrance to the labor market into account. On the other hand, it should also include social costs and benefits of e.g. teacher costs from the additional year in kindergarten / public schooling, the costs from one year later entrance to the labor market with respect to lost tax revenue, the social benefits of lower crime rates, etc. Yet, the PRTEs provide the foundation for future cost/benefit analyses of different policies directed towards school starting age.

VI. Conclusion

This paper uses Danish register-based data to investigate the effect of school starting age on crime while using exogenous variation in school starting age generated by administrative rules.

We find that a higher school starting age lowers the propensity to commit crime before age 18. For boys property crime is reduced while for girls violent crime is reduced. Importantly, we find that the boys who benefit the most from higher school starting age are those with high levels of latent abilities whereas those with low levels of latent ability are unaffected by school starting age. In

contrast, for girls the effects are homogeneous across levels of latent ability. We also find that the effects are not caused by relative age of peers but by one's own school starting age.

Detailed studies of the age-profile of the effects indicate that the reductions to crime are likely to be caused by an incapacitating effect of schooling, as those who start school later graduate later. Although not directly testable, the pattern of results supports this hypothesis: Boys who are old-for-grade are less likely to be charged during the period until they turn 17 years, and this effect stems from crime (primarily property crime) taking place during the weekdays. Girls who are old-for-grade are more likely to postpone the debut of a criminal career until after age 19, and this effect stems from crime (mainly violent crimes) taking place Friday night. Only for boys, we find small effects on highest attained education as measured at age 27.

Finally, we estimate the effects of different policy changes. Increasing school starting age could lower crime – more so for boys than for girls. Yet, our findings do not suggest that school starting age should be increased. Postponing school entrance is costly and we find that higher school starting age only reduces crime through incapacitation while boys with low level of latent ability - who often is the intended target group when discussing school starting age - are unaffected by higher school starting age. Scarce resources may in consequence be better spent if they are redirected from postponing school start to other interventions, as there is ample evidence of the benefits of early childhood interventions on e.g. crime and education.

Literature

Anderson, M. (2012): In School and Out of Trouble? The Minimum Dropout Age and Juvenile Crime. Accepted conditional upon minor revisions at *Review of Economics and Statistics.*

Björklund, A. & R. Moffitt, (1987): The Estimation of Wage Gains and Welfare Gains in Self-selection *Review of Economics and Statistics*, 69(1), 42-49

Black, S. E., P. J. Devereux and K. G. Salvanes (2007): From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. *Quarterly Journal of Economics,* 122(1), 409-439.

Black, S. E., P. J. Devereux and K. G. Salvanes (2011): Too young to leave the nest? The effects of school starting age. *Review of Economics and Statistics* 93, 455-467.

Black, S. E., P. J. Devereux and K. G. Salvanes (2013): Under Pressure? The Effect of Peers on Outcomes of Young Adults. *Journal of Labor Economics* 31(1):119-153.

Crawford, C., L. Dearden and C. Meghir (2010): When you are born matters: the impact of date of birth on educational outcomes in England. DoQSS WP 10-09.

Cunha, F., J. Heckman (2008): Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, 43(4), 738-782

Dalsgaard, S., M. K. Humlum, H. S. Nielsen and M. Simonsen (2012): Relative standards in ADHD Diagnoses: The role of specialist behavior. *Economics Letters* 117, 663-665.

Elder, T. E., (2010): The importance of relative standards in ADHD diagnoses: Evidence based on exact birth dates. *Journal of Health Economics* 29, 641-656.

Evans, W. N., Morrill, M. S., and S. T. Parente (2010): Measuring excess medical diagnosis and treatment in survey data: the case of ADHD among school-age children. *Journal of Health Economics* 29, 657-673.

Fredriksson, P. and B. Öckert (2005): Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance *IZA DP* #1659.

Gaviria, A., S. Raphael (2001): School Based Peer Effects and Juvenile Behavior, *Review of Economics and Statistics*, 83, 257-268.

Heckman, J., S. Urzua and E. Vytlacil, (2006):. Understanding Instrumental Variables in Models with Essential Heterogeneity. *Review of Economics and Statistics*, 88(3), 389-432

Heckman, J., E. Vytlacil, (2005): "Structural Equations, Treatment Effects, and Econometric Policy Evaluation" *Econometrica*, 73(3), 669-738

Jacob, B. and L. Lefgren (2003): Are Idle Hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime. *American Economic Review* 93 (5), 1560-1577.

Lochner, L. (2011), Non-Production Benefits of Education: Crime, Health, and Good Citizenship. in Hanushek, Machin and Woessman, eds., Handbook of Economics of Education, vol.4. Elsevier

Luallen, J. (2006) School's Out...Forever: A Study of Juvenile Crime, At-Risk Youths and Teacher Strikes. *Journal of Urban Economics*, 59:75-103.

OECD (2009): *Society at a Glance 2009. OECD Social Indicators*. OECD, Paris, France.

Sacerdote, B. (2001): Peer Effects With Random Assignment: Results from Dartmouth Roommates. *Quarterly Journal of Economics*, 116, 681-704.

Appendix A

TABLE A1

FRACTION OF STUDENTS BEING RETAINED AT EACH GRADE LEVEL

Grade level	Fraction delayed/retained
Kindergarten	0.136
1st grade	0.014
2nd grade	0.003
3rd grade	0.004
4th grade	0.003
5th grade	0.003
6th grade	0.003
7th grade	0.002
8th grade	0.003
9th grade	0.005

Note: Calculations based on grade levels from 2007 and onwards.

TABLE A2

SUMMARY STATISTICS OF THE SAMPLE

Variable	Girls			Boys		
	December	January	Difference	December	January	Difference
Immigrant	0.043	0.037	0.006**	0.042	0.034	0.008***
	0.001	0.001		0.001	0.001	
Parents married/cohabiting	0.788	0.784	0.004	0.789	0.792	-0.003
	0.003	0.003		0.003	0.003	
Apgar score=9	0.181	0.184	-0.002	0.187	0.184	0.004
	0.002	0.002		0.002	0.002	
Apgar score=8	0.071	0.067	0.004	0.066	0.070	-0.004
	0.002	0.002		0.002	0.002	
Apgar score lower	0.084	0.085	-0.001	0.097	0.094	0.003
	0.002	0.002		0.002	0.002	
Birth weight, grams	3341	3358	-16.59**	3466	3481	-15.02**
	3.987	3.813		4.152	3.909	
Gestational length, weeks	39.564	39.543	0.021	39.482	39.464	0.019
	0.012	0.011		0.012	0.012	
Mother:						
- Months of schooling	137.276	137.502	-0.226	136.986	138.480	-1.494***
	0.230	0.235		0.229	0.226	
- Completed HS or equvalent	0.287	0.230	0.057	0.290	0.302	-0.013**
	0.003	0.003		0.003	0.003	
- Unemployed	0.130	0.125	0.005	0.128	0.122	0.006*
	0.002	0.002		0.002	0.002	
- Out of the labor force	0.104	0.111	-0.006*	0.105	0.105	0.000
	0.002	0.002		0.002	0.002	
- Age at birth of first child	24.819	24.886	-0.068	24.851	24.990	-0.139***
	0.027	0.027		0.026	0.026	
Father:						
- Months of schooling	139.736	140.507	-0.771	139.544	141.215	-1.671***
	0.281	0.277		0.277	0.273	
- Completed HS or equvalent	0.191	0.190	0.002	0.195	0.204	-0.009*
	0.003	0.003		0.002	0.003	
- Unemployed	0.078	0.077	0.001	0.076	0.073	0.004
	0.002	0.002		0.002	0.002	
- Out of the labor force	0.065	0.062	0.003	0.002	0.002	0.000
	0.002	0.002		0.064	0.001	
Number of observations	24279	24267		25157	25226	

Note: * indicates $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

TABLE A3

EFFECT OF SCHOOL STARTING AGE ON CRIME ACROSS THE WEEK

	Girls	Boys
Crime prior to 18th birthday		
Monday	-0.005 *	0.002
	-2.21	0.31
Tuesday	0.001	-0.013 *
	0.24	-2.49
Wednesday	-0.002	-0.006
	-0.91	-1.21
Thursday	0.000	-0.004
	0.21	-0.81
Friday	-0.005 *	-0.008
	-2.33	-1.44
Saturday	-0.001	0.000
	-0.64	0.03
Sunday	-0.002	-0.005
	-1.66	-1.14
# Observations	48,546	50,383

Note: * indicates $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

FIGURE A1

AREA OF COMMON SUPPORT, GIRLS

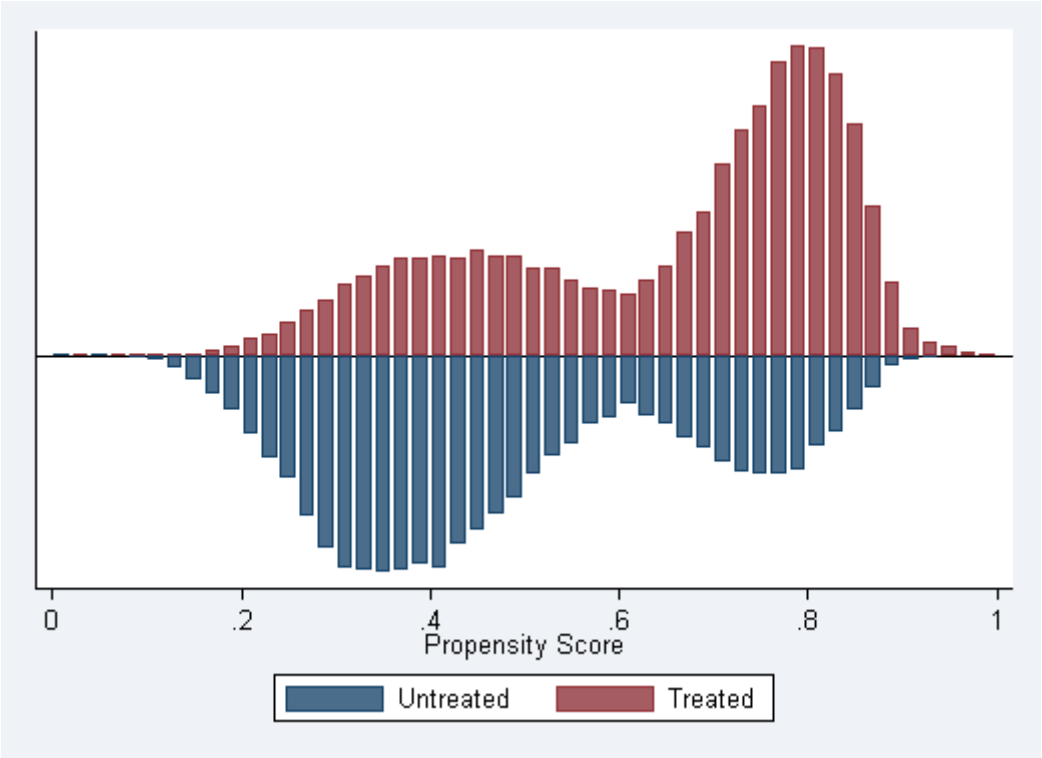


FIGURE A2

AREA OF COMMON SUPPORT, BOYS

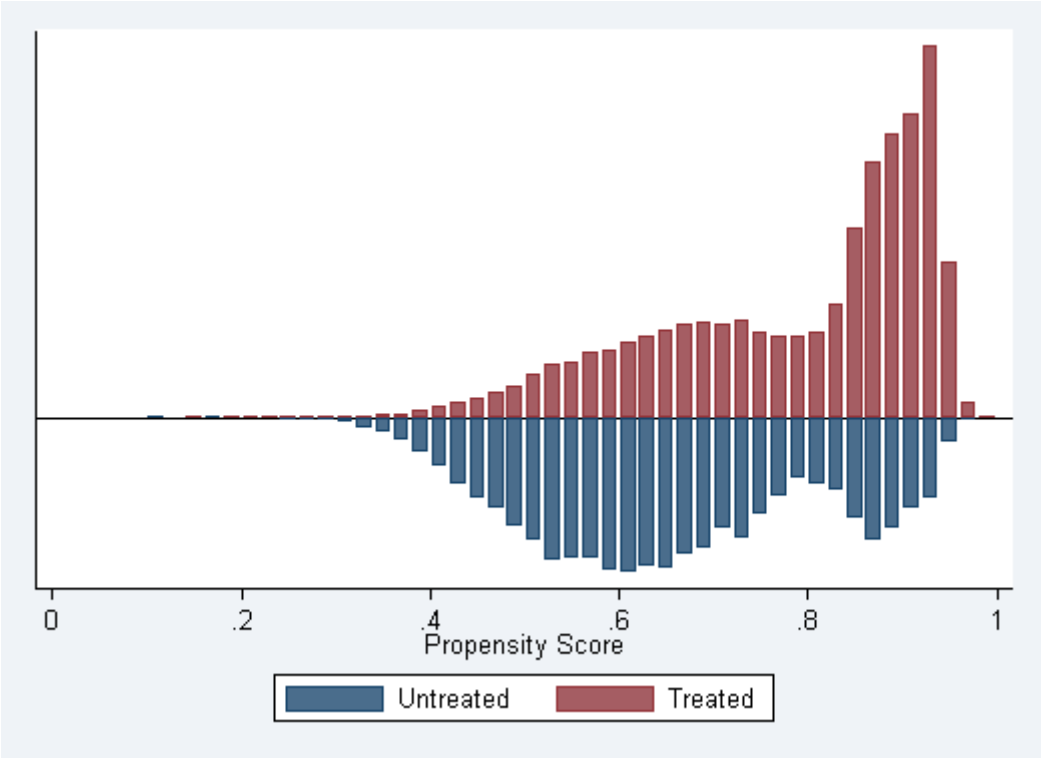


FIGURE A3

POLICY RELEVANT TREATMENT EFFECTS WEIGHTS, GIRLS

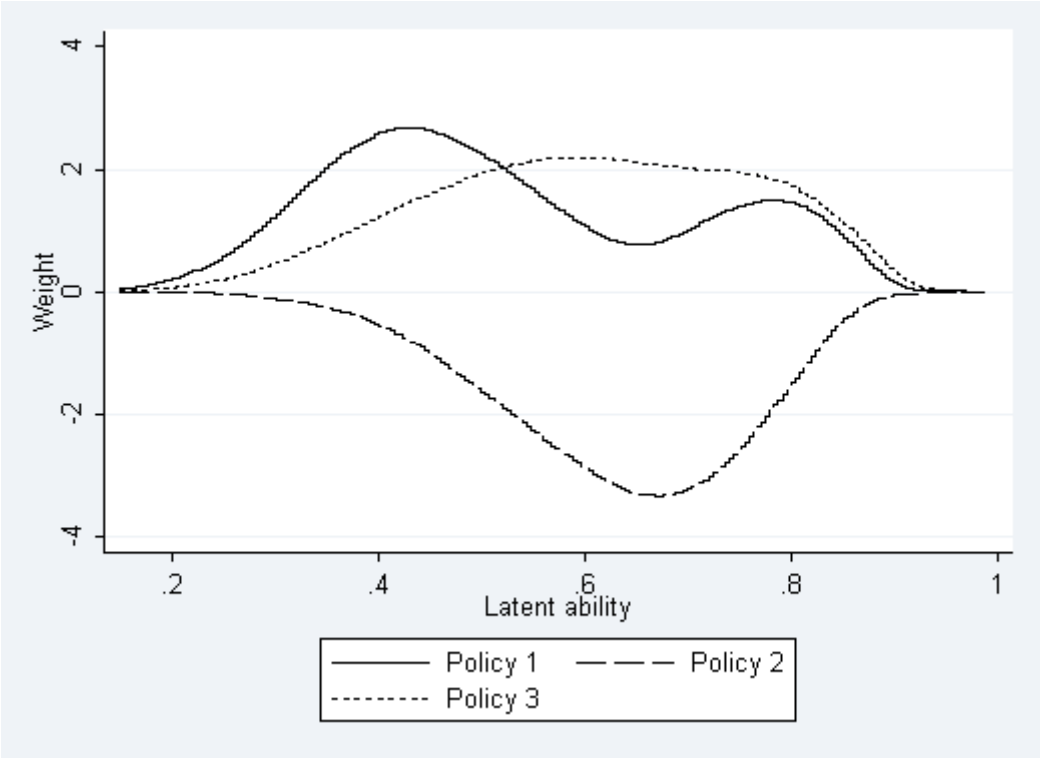


FIGURE A4

POLICY RELEVANT TREATMENT EFFECTS WEIGHTS, BOYS

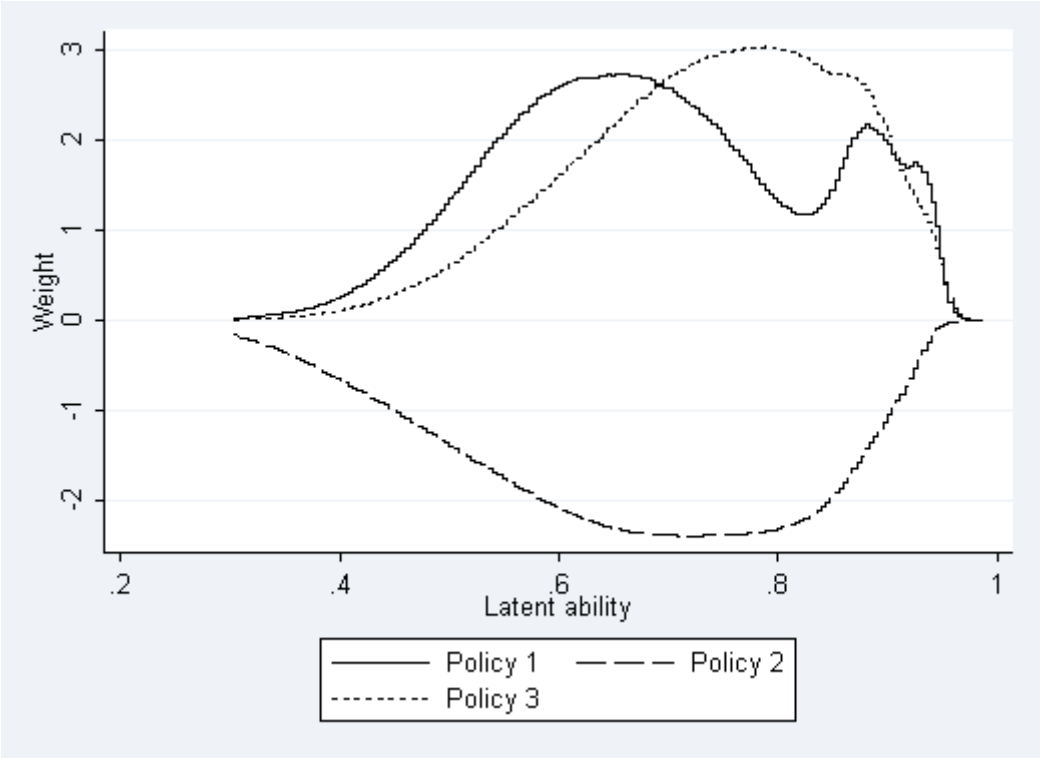
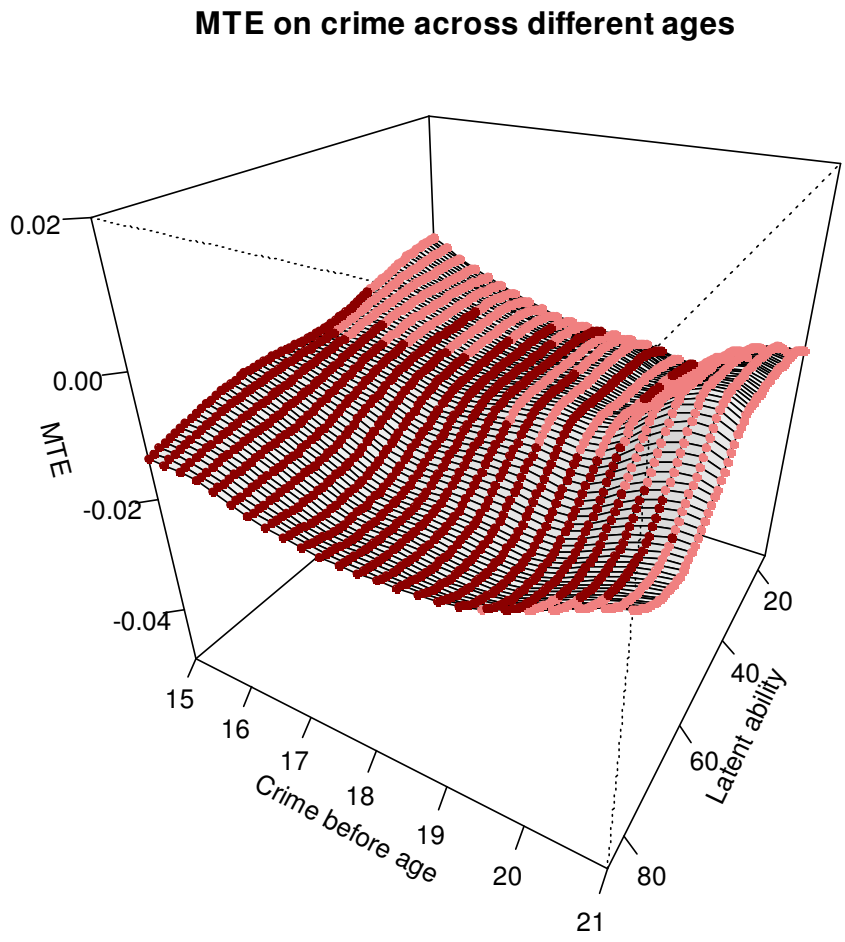


FIGURE A5

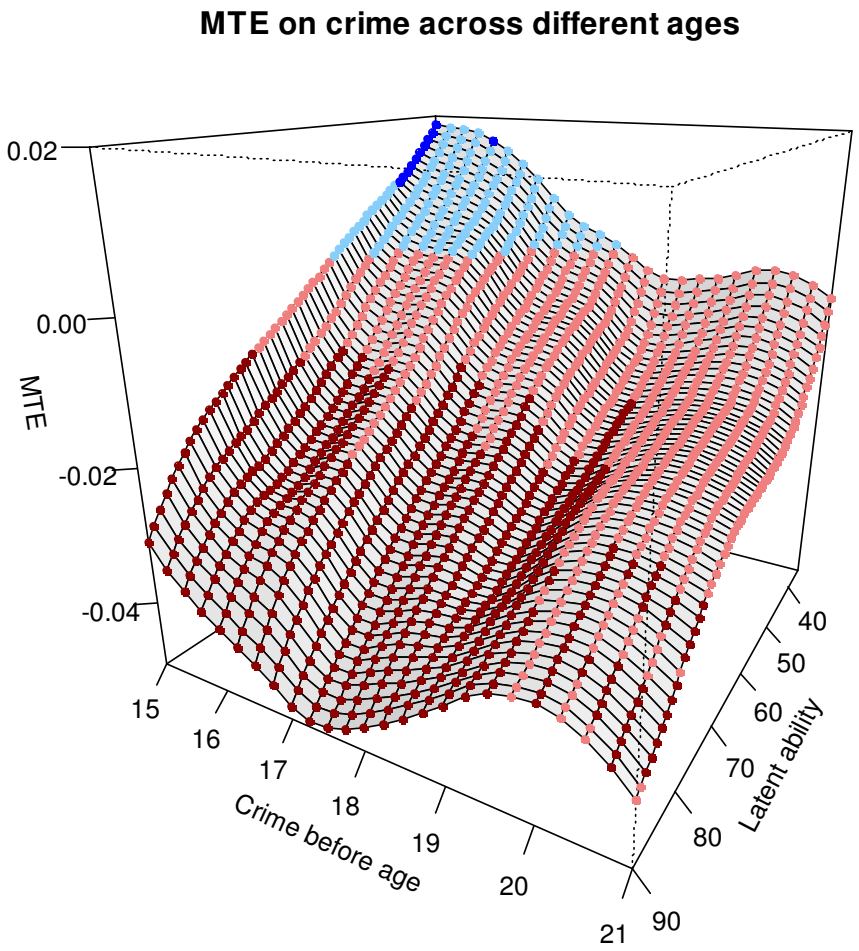
MARGINAL TREATMENT EFFECTS, CRIME ACROSS DIFFERENT AGES, GIRLS



Note: Red indicates significant and negative estimates, light red indicates negative and insignificant estimates, light blue indicates positive and insignificant estimates, and blue indicates positive and significant estimates.

FIGURE A6

MARGINAL TREATMENT EFFECTS, CRIME ACROSS DIFFERENT AGES, BOYS



Note: Red indicates significant and negative estimates, light red indicates negative and insignificant estimates, light blue indicates positive and insignificant estimates, and blue indicates positive and significant estimates.

Appendix B (not to be published)

Table B1 shows the results from replication of BDS (2011). The OLS results indicate that children who are old-for-grade and start school at age 7.6 are significantly less likely to enroll in upper secondary school by age 18/two years after completing 9th grade. To be able to compare our results with Black et al. (2011), we also show effects on length of completed schooling (by age 27) for a subset of the cohorts under analysis. Again, the OLS results indicate that children with a relatively high school starting age perform much worse than those who start earlier, in line with the Norwegian analysis. The IV results, on the other hand show a different pattern; in fact, all children are significantly *more* likely to start secondary school early. For boys we also find a small significant effect on months of completed schooling of about four months of additional schooling. Our effects on completed education are slightly larger than those reported by Black et al. (2011), which might be explained by the different set of cohorts (1962-1988 vs. 1981-1988).With regards to teenage pregnancies, OLS results indicate that being old-for-grade is associated with worse outcomes: girls are significantly more likely to have given birth at age 19/two years after completing 9th grade. Yet again, this seems to be caused by endogeneity. When instrumenting old-for-grade, girls are actually almost a percentage point *less* likely to have given birth at age 19 if they start school at age 7.6. This is in line with the findings from Black et al. (2011). There is little effect on the probability of being a teenage mother two years after completing 9th grade.

Figure B1 and figure B2 show the MTEs of being old-for-grade on enrollment into post-compulsory school by the age of 17 for girls and boys respectively. Post-compulsory school is a dummy variable, so an estimate of e.g. 0.01 implies that increasing school starting age by one year results in an increase to the probability of enrollment at the age of 17 by 1 %-point.

TABLE B1

REPLICATION OF BLACK, DEVEREUX AND SALVANES (2011)

Variable	Girls						Boys					
	OLS		2SLS		2SLS		OLS		2SLS		2SLS	
Enrolment (0/1)												
- in upper secondary/vocational school by age 18	-0.048	***	0.094	***	0.090	***	-0.014	**	0.177	***	0.160	***
	-13.61		9.99		9.96		-3.44		13.5		12.66	
- in upper secondary/vocational school 2 years after 9th	-0.060	***	0.062	***	0.058	***	-0.021	***	0.138	***	0.121	***
	-16.97		6.53		6.39		-5.12		10.42		9.49	
# Observations	48,546		48,546		48,546		50,383		50,383		50,383	
Months of schooling (by age 27, born 1985 or earlier)	-7.638	***	-0.010		-0.314		-4.882	***	3.946	**	2.393	*
	-17.16		-0.01		-0.31		-10.33		3.03		1.99	
# Observations	14,169		14,169		14,169		14,338		14,338		14,338	
Teenage pregnancy (0/1)												
- by age 19	0.003	*	-0.009	**	-0.008	**						
	2.45		-3.18		-3.14							
- 2 years after 9th grade	0.008	***	-0.001		-0.001							
	9.17		-0.47		-0.41							
# Observations	43,668		43,668		43,668							
Controls												
- *Yearly cut-off fixed effects*	X		X		X		X		X		X	
- *Covariates*					X						X	

Note: * indicates $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

FIGURE B1

MARGINAL TREATMENT EFFECTS, ENROLLMENT BY AGE 18, GIRLS

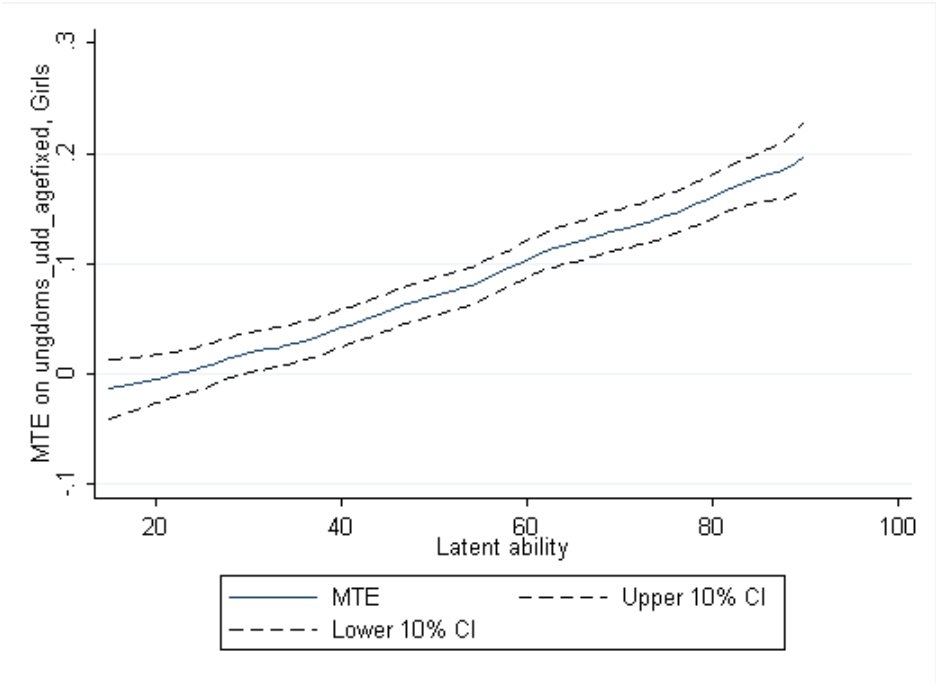
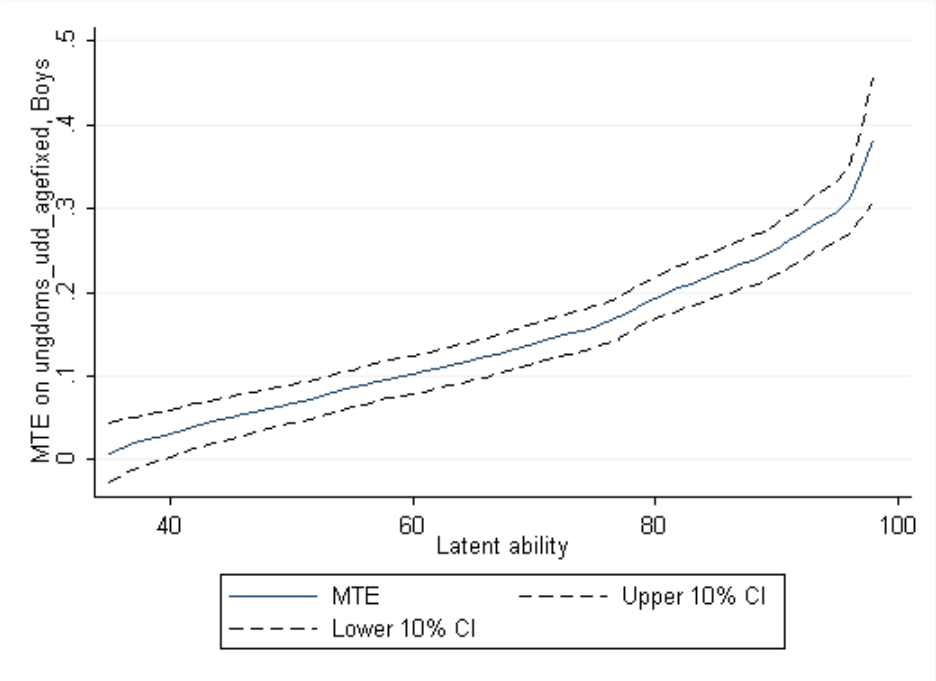


FIGURE B2

MARGINAL TREATMENT EFFECTS, ENROLLMENT BY AGE 18, BOYS



We see from figure B1 that the estimated effects of being old-for-grade on enrollment into post-compulsory schooling for girls are close to zero and insignificant at the lowest levels of latent ability. However, the estimated effects increase in latent ability, to almost 0.2 for the most able girls. From figure B2 we see a similar relationship for boys. At the lowest intervals of latent ability, the marginal treatment effects of being old-for-grade are close to zero and insignificant. As latent ability increases so does the estimated effects to around 0.3 for the most able boys. The averages of the two MTEs are approximately 0.1 for girls and 0.15 for boys. These correspond to the 2SLS estimates from table B1.

Figure B3 and figure B4 show the MTEs of being old-for-grade on months of schooling by the age of 27 for girls and boys respectively. An estimate of e.g. 3 implies increasing school starting age by one year results in an increase of total educational attainment at age 27 by 3 months.

FIGURE B3

MARGINAL TREATMENT EFFECTS, MONTHS OF SCHOOLING BY AGE 27, GIRLS

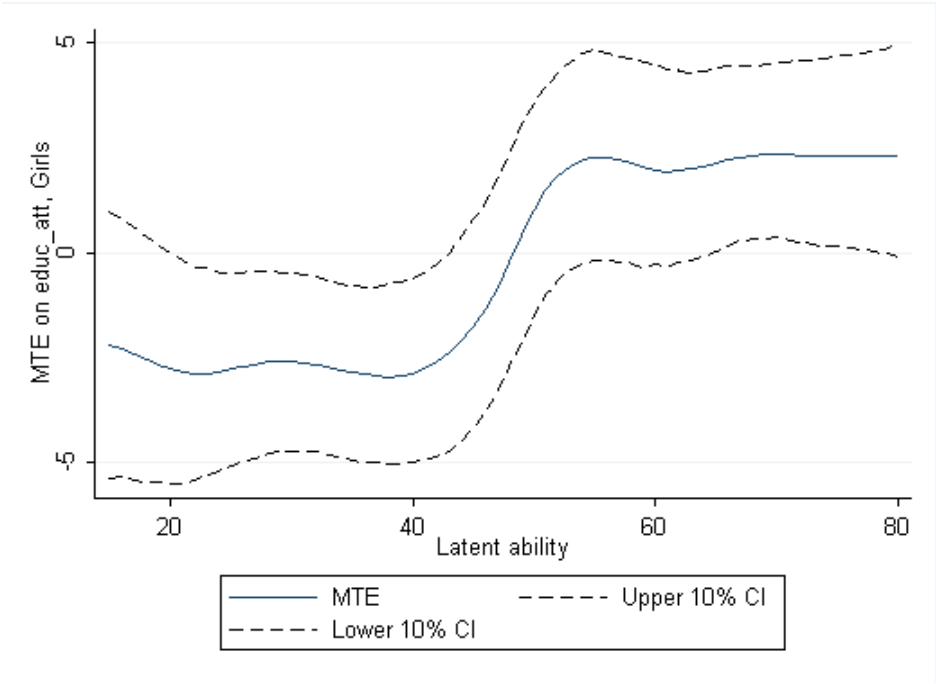
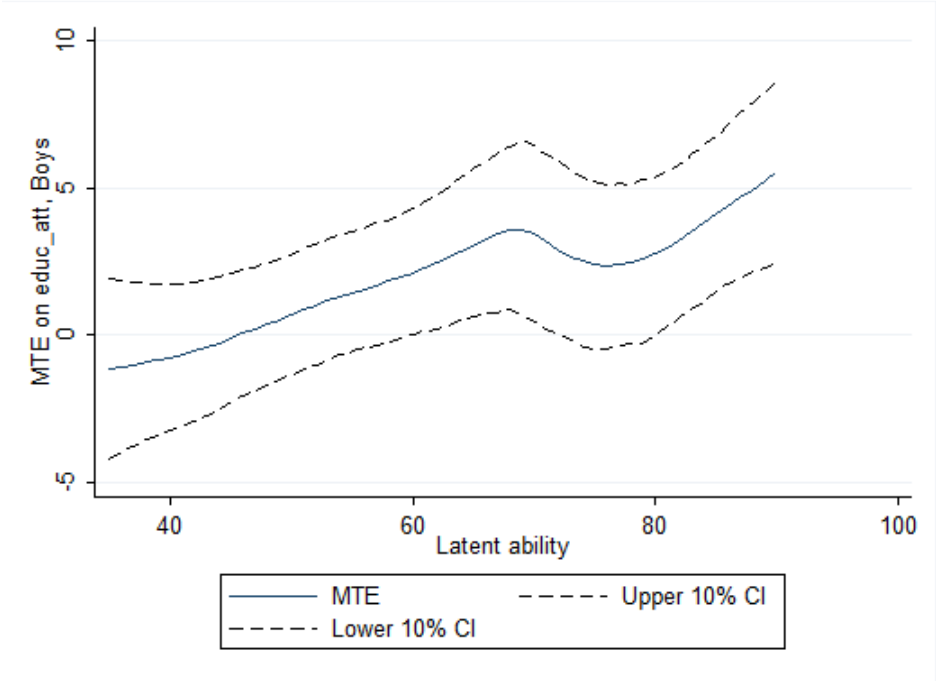


FIGURE B4

MARGINAL TREATMENT EFFECTS, MONTHS OF SCHOOLING BY AGE 27, BOYS

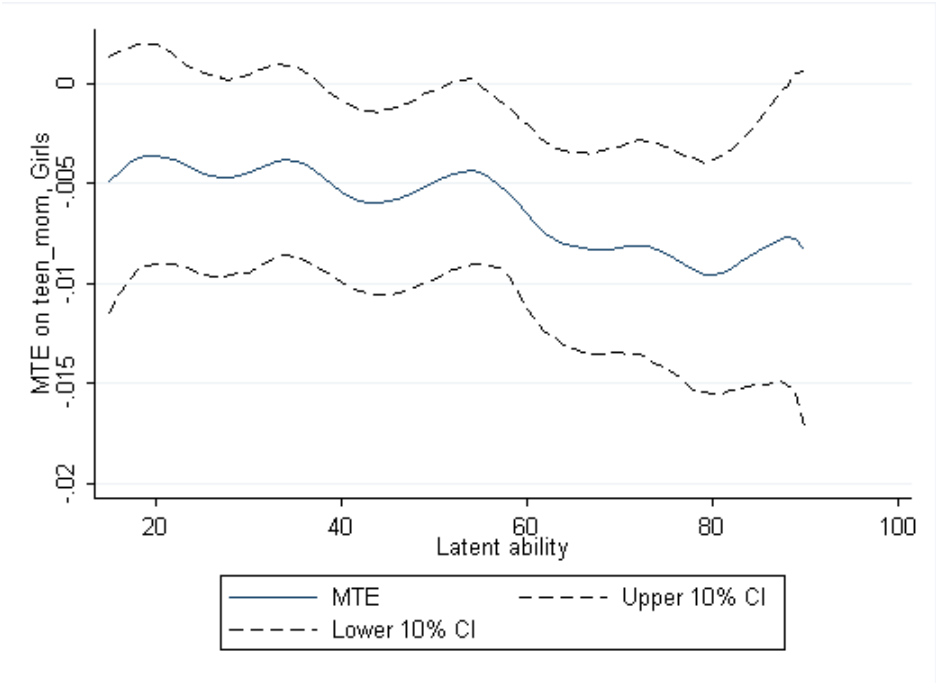


We see from figure B3 that the estimated marginal treatment effects being old-for-grade on total months of education at age 27 are negative and borderline significant at the 10 % level for the least able girls (i.e. those with low values of latent ability). The figure shows that for the least able girls being old-for-grade is estimated to reduce total schooling by 2 to 3 months. As with the previous figures, the estimated effects increase in latent ability, so the estimated effects for the most able girls are positive and borderline significant at the 10 % level, with estimated effects around 3 months. The 2SLS from table B1 showed that the average effect of being old-for-grade for girls on months of schooling were insignificant and close to zero. The 2SLS estimate for boys were 2.5. Again we see that the MTEs reveal some heterogeneity. From figure B4 we that the least able boys appear to be unaffected by being old-for-grade, whereas the most able boys appear to benefit somewhat from having a higher school starting age with estimates around 3-4 months.

Figure B5 shows the MTEs of being old-for-grade on risk of teenage pregnancies prior to the age of 19 for girls. Teenage pregnancies is a dummy variable, so an estimate of e.g. -0.01 implies that increasing school starting age by one year results in a reduction to the probability of having given birth to a child at the age of 19 by 1 %-point.

FIGURE B5

MARGINAL TREATMENT EFFECTS, TEENAGE PREGNANCY AGE 19, GIRLS



We see from figure B5 that being old-for-grade is estimated to have either an insignificant or negative impact on girls' probability of having a child prior to their 19th birthday, depending on latent ability. The estimated marginal treatment effects are insignificant at the lower levels of latent ability and significant with estimates around -0.01 at the higher levels of latent ability. As with the previous figures, figure B5 shows that the average of the MTEs reflects the 2SLS from table B1 (-0.007).