School Starting Age and the Crime-Age Profile

Rasmus Landersø
Rockwool Foundation Research Unit & Department of Economics and Business, Aarhus University

Helena Skyt Nielsen
Department of Economics and Business
Aarhus University

Marianne Simonsen
Department of Economics and Business
Aarhus University

June 2014

Abstract:

This paper uses register-based data to investigate the effects of school starting age on crime and the determinants of crime-age profiles. We exploit that Danish children typically start first grade in the calendar year they turn seven, which gives rise to a discontinuity in school starting age for children born around New Year. We find that higher school starting age lowers the propensity to commit crime at young ages and that this to some extent is driven by incapacitation. We also find persistent effects on the number of crimes committed for boys. Finally, individuals with high latent characteristics benefit most.

JEL: I21, K42.

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

Acknowledgements: Comments from Matthew Lindquist, Naci Mocan, Magne Mogstad as well as seminar participants at BeNA Humboldt University, ISER, University of Essex, University of Sussex, KORA, University of St. Gallen, Xiamen University, the joint University of Bergen & UCL workshop 2013, the workshop on "Economics of Successful Children" at Aarhus University 2013, the University of Chicago's Life Cycle and Inequality working group as well as participants at ESPE 2013, EALE 2013, ASSA 2014 and at the DGPE workshop are appreciated. Financial support from the Danish Council for Strategic Research (CSER, 09-070295) and CIRRAU (Simonsen) is gratefully acknowledged. 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 (e.g. Bedard and Dhuey (2006)). Black, Devereux and Salvanes (2011) and Crawford, Dearden and Meghir (2010) refine this type of analysis and show that this 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. We also provide novel insights into the determinants of criminal behavior more generally, as we use the mechanical relationship between delayed school entry and delayed life-course to address whether the observed crime-age relationship is caused by age or by the timing of key life experiences. In addition, we investigate heterogeneity in effects across different types of crimes, timing of crime throughout the week, across ages and across the distribution of latent characteristics.

As in Black et al. we rely on exogenous variation in 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 key 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 at a given age until age 20 where the effects begin to fade out. In addition, boys experience a persistent reduction in the *number* of crimes committed. We use these results to show that the crime-age profile is largely caused by life course rather than age *per se*. We then extend our analysis and show that the LATE estimates of school starting age using an administrative cut-off date cannot be generalized to the population in general as effects differ with the likelihood that a child has a higher school starting age. By estimating marginal treatment effects we find evidence of essential heterogeneity for boys but not for girls. We find that boys with high levels of latent characteristics experience the largest benefits in terms of crime reductions and school enrollment, whereas boys with intermediate levels of latent characteristics seem to be unaffected by starting later in school. Furthermore, we investigate potential mechanisms behind the crime reductions during the teen-years and find that incapacitation seems to play an important role. For boys, being old-for-grade reduces criminal charges until age 19, and the effect is driven by property crime and crime committed on weekdays. For girls, being old-for-grade postpones initiation of crime, and the effect is driven by violent crime and crime committed on Fridays. Finally, we find that the relative age of classroom peers does not seem to be behind the reduction in crime.

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, 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 ours because there is no automatic relationship between school starting age and minimum required schooling as there would be in the US and 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).[2]

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.[3]

Obviously, school starting age is not random and is most likely affected by a range of factors that may also correlate with the child's outcomes, behavioral as well as academic. For example children's overall school readiness and behavior in preschool is likely to affect the timing of school start. This pattern is clear from Figure 1 that shows the distributions of social and emotional difficulties at age 4 among punctual and late school starters, drawn from an auxiliary data source

[1] The school starting regulations are not strictly enforced and exemptions are granted based on applications from the parents. For details consult the Danish Law of public schools.

[2] It was also possible to complete a voluntary tenth grade before continuing on to a vocational or academic path.

[3] A minimum of 67 % of the expenses is covered by the local authorities, c.f. the Danish Law of day care.

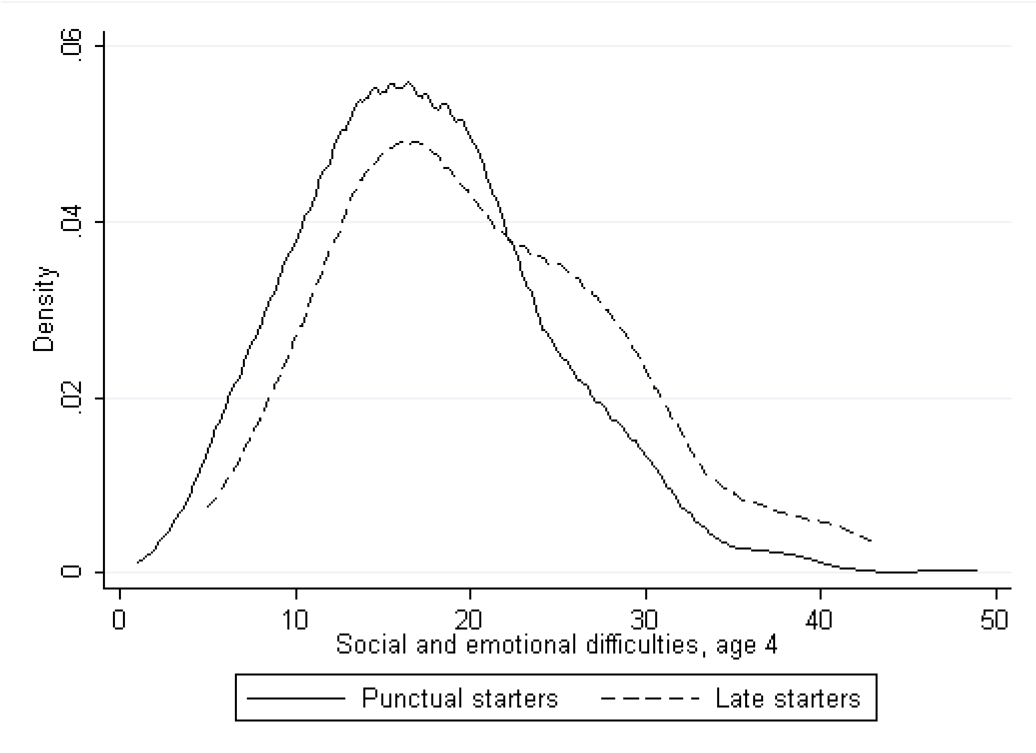
(the Danish Longitudinal Survey of Children).[4] But other factors may impact on the decision as well: as shown by the previous literature, starting later is likely to increase test scores, and while this has not been found to impact significantly on long term outcomes such as earnings, higher grades may improve the consumption value of attending school. Higher grades also allow for a more extensive educational choice set thus securing a better match between student preferences and type of degree, potentially allowing students to choose more prestigious degrees. Finally, there is considerable variation in school starting age culture across municipalities even conditional on a rich set of observable characteristics. For completeness, we will investigate some of these hypotheses towards the end of the paper.

[4] 'Punctual school starters' obey the rules and start school when they are supposed to start according to the rules, while 'late school starters' have been granted an exemption.

FIGURE 1

AGE 4 SOCIAL AND EMOTIONAL DIFFICULTIES

AMONG PUNCTUAL AND LATE SCHOOL STARTERS



Note: Data stem from the Danish Longitudinal Survey of Children that surveys children born in September and October of 1995.

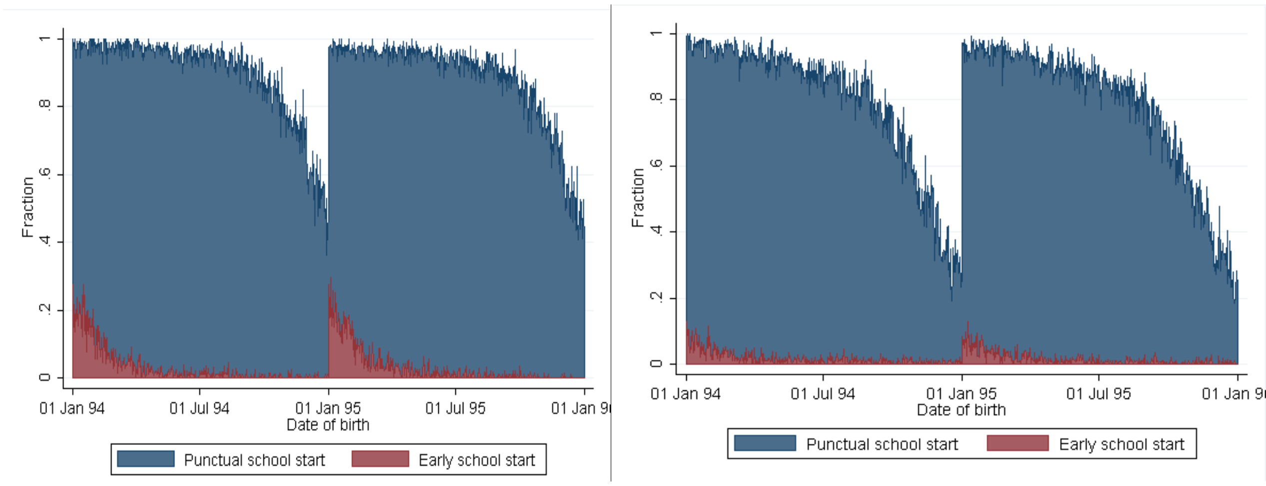
To meaningfully address consequences of school starting age, our empirical analysis will make use of the following observation: because the formal age at school start is defined by the year of birth, each January 1st provides a cut-off point at which children born on each side are subject to a one year difference in timing of administratively determined school start, even though they are born very close in time. Section III will formalize this idea. 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.[5] 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. As shown in Figure 2, 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. If children born around the cut-off date are 7.6 years old at school start, we label them "old-for-grade". Figure 3 shows the fraction of children who are old-for-grade by date of birth for each gender.

FIGURE 2

FRACTION PUNCTUAL, EARLY, AND LATE SCHOOL START,

BY DATE OF BIRTH. SELECTED COHORTS

GIRLS BOYS

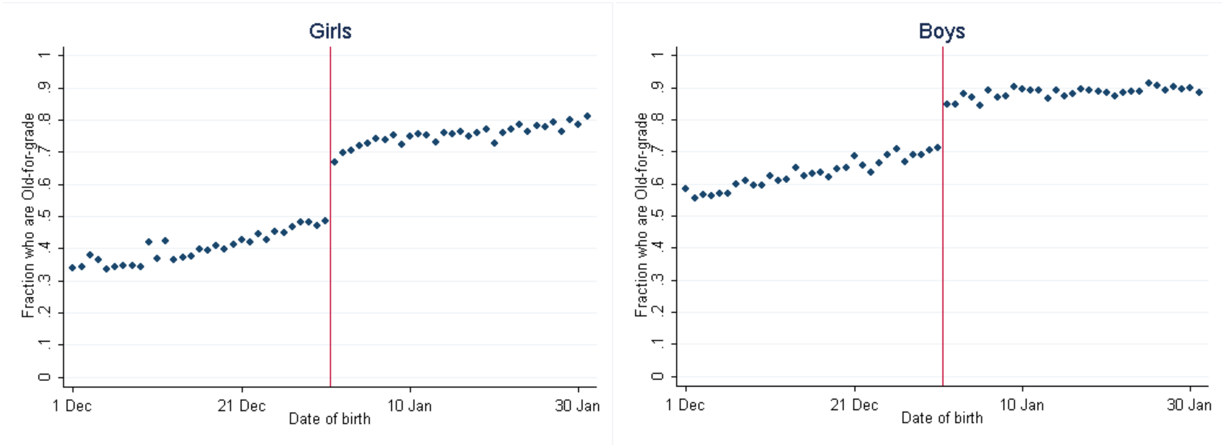


Note: Population of children born January 1 1994 – January 1 1995.

[5] 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 3

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

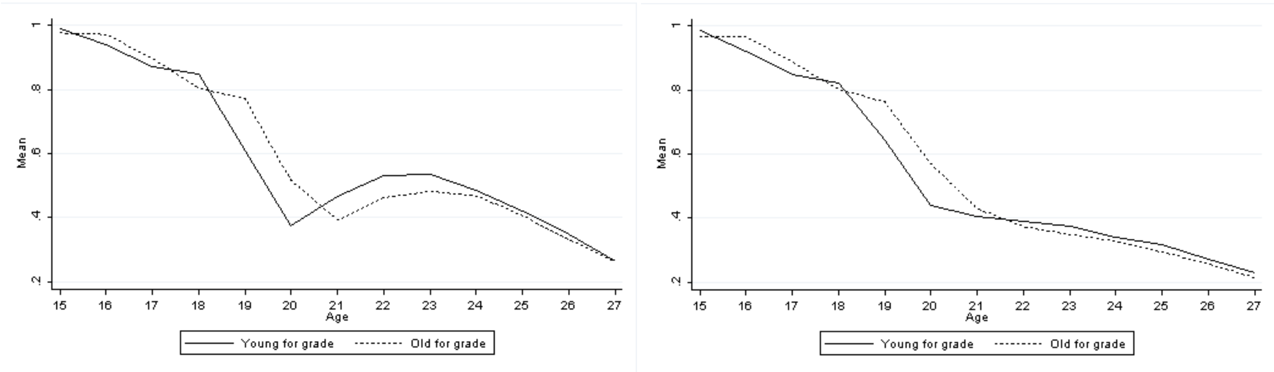


Note: Being old-for-grade implies that the child starts school at age 7.6 instead of at age 6.6. Average for population of children born December 1981 to January 1993.

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 much more likely than girls to be old-for-grade.

The crime-age profile refers to an almost universally observed relationship between crime rates and age, where crime rates increase until the peak around age 18-20 and then decrease for the remainder of the life (Gottfredson and Hirschi, 1990). School starting age may affect the crime-age profile because it changes the timing of key lifetime experiences. In our setting, delaying school start by one year will, for a given education length, also delay graduation and labor market entry by one year. Figure 4 illustrates this by showing the average fraction of old-for-graders and young-for-graders who are enrolled into education.

FIGURE 4: Fraction of girls (left) and (boys) enrolled in education by age



The question is whether criminal behavior is tied to one's actual age (see Gottfredson and Hirschi, 1990) or tied to turning points in life such as for instance school leaving and educational attainment (Sampson and Laub, 1995).

One reason to anticipate an effect of enrollment on crime is incapacitation: when youth are in school, they simply have less time to commit crime (see Lochner, 2011). 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 but 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). Cook and Kang (2014) find that higher age at school start increases crime due to higher exposure to dropout in the US institutional setting with minimum dropout age rather than a certain number of compulsory school grades. Anderson (forthcoming) finds that increasing the compulsory school *leaving* age reduces arrests for both violent and property crime. Anderson (forthcoming) estimates that a minimum dropout age of 18 decreases arrest rates for 16-18 year-olds by 17 %. This effect may be due to incapacitation. An alternative explanation might be that 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.

Another mechanism by which school starting age may affect later behavior is 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 other outcomes such as the tendency to engage in criminal activities.

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.

A fourth mechanism by which school starting age (SSA) likely affects crime is through changing the opportunity costs of crime (see Grogger (1998)), because potential wage earnings are closely linked to educational attainment. As we show in this paper, postponing school start delays graduation and thus affects the opportunity costs of crime at a given age.

We formally investigate the importance of the postponement of school leaving and in particular the importance of incapacitation effects. In addition we check if our results are robust to including 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.[6] 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. This is true regardless of the severity of the crime, and there is no such thing as a youth court.[7] 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.[8] 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.

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

[7] The question of quilt is, in fact, never determined for children below the age of criminal responsibility. The severity of the case is solely considered by the Attorney General.

[8] See the Danish Service Act.

Reported victimization rates in Denmark are falling like in the rest of the OECD. However, overall crime rates in Denmark are somewhat higher than in 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.

III. Methodology

Our goal is to estimate the effect of changing school starting age from young-for-grade (SSA_0) to old-for-grade (SSA_1) on the associated crime outcomes:

(1) $\Delta = Y_1 - Y_0$

(2) $Y_j = M_j(X) + U_j, j = 0,1$

where Y_j denote the potential outcome associated with SSA_j , X observable characteristics[9], and U_j unobservable characteristics. U_1 and U_0 are likely related to the choice of school starting age which would bias results if ignored. To circumvent the problem that *SSA* is not random and likely related to *Crime* through U_j, we formally 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 before 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

[9] X 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-off, and a constant.

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.

As illustrated above, 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:

(3) $$OG = 1[\mu(X, Z) - V > 0]$$

Where Z is an indicator for being born before the cut-off date at January 1st: $Z = 1[birth = December]$. Employing a linear probability model, we get the following first stage equation:(4)

$$OG = X\gamma_1 + \gamma_2 1[birth = December] + \tilde{V}$$

The indicator for being born in December (z) is plausibly excludable from the main equation of interest (crime). Moreover, monotonicity is fulfilled if we assume that no one starts school at earlier ages if born in January instead of December. This is not a strong assumption. Put differently, a defier in our set-up is a child whose parents choose to enroll him *earlier* (at age 6.6) than at the age specified by administrative rules if born in January and *later* (at age 7.6) than at the age specified by administrative rules if born in December. With a relevant and excludable instrument and an assumption of monotonicity, we can estimate the Local Average Treatment Effect of *OG* as[10]:

(5) $$\beta_2^{LATE} = E\{\beta_2 | X\gamma_1 < V < X\gamma_1 + \gamma_2\}$$

[10] In most of our empirical results presented below, we will use the year-to-year variation of the first stage effect of the cut-off to gain additional support. Doing so implies the assumption that the year-to-year variation in our first stage results is independent of the year-to-year variation in crime rates. Our results are robust to disregarding the year-to-year variation so we only have one IV.

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:

(6) $$\beta_2^{MTE} = E\beta_2 | U_D = u_D^* \quad , X = X^*$$

where $U_D = F_V[V]$.[11] 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 U_D. Following the selection equation (equation 3), the margin of treatment is the point where treatment probability by observed characteristics $\mu(X,Z)$ must equal latent characteristics U_D (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_D (V) to offset the disadvantageous observable characteristics and make $\mu(X,Z) = V$. Oppositely, individuals on the margin of treatment with low observed treatment probabilities must have small values of U_D. $U_D = F_V[V]$ can therefore be interpreted as a scale measuring increasing levels of latent abilities.[12]

[11] 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.

[12] We use the words latent characteristics and latent abilities interchangeably, although it is inherently unobservable. However, the main factors predicting whether a child is old-for-grade or not (other than birth date) are unfavorable background characteristics such as low birth weight or parental education. See Section V.A.

Hence, the MTEs allow us to distinguish between heterogeneous treatment effects across latent characteristics U_D.[13]

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[14] 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[15], demographic variables, educational variables, and crime by the unique individual identification number. We can link parents 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, 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.[16]

[13] We estimate the Marginal treatment effects using a semi-parametric setup assuming separability as outlined in Brinch, Mogstad and Wiswall (2014).

[14] Using charges instead of convictions enables us to use three additional years of data because of the right to appeal. Conclusions are robust to using convictions instead of charges.

[15] The APGAR score ranges from zero to 10 and summarizes the health of a newborn child based on five simple criteria: Appearance, Pulse, Grimace, Activity and Respiration.

[16] 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.

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 and Figures A1 and A2). 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.[17] In addition, there is no relationship between the cutoff and being redshirted or skipping grades during primary school.

Table A2 in the 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.[18] We include these variables as covariates and experiment with further restrictions on the birth date intervals, and our results are robust to these restrictions.

As is true for most of the existing literature on school starting age, choosing the right outcome is a challenge: on the one hand, one wants to align children in terms of age. This is particularly relevant because crime is positively correlated with age in the age range considered in this paper. On the other hand, one wants to align children in terms of length of education because the agents that decide a child's SSA may focus on these outcomes or because education may have a direct effect on

---

[17] Measurement errors in school starting age will impact on our results to the extent that they are correlated with the instrument. If children born in December (z=1) are more likely to repeat a grade as suggested by Elder and Lubotsky (2009), our results will be biased towards zero.

[18] 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).

the tendency to commit crime. To address these issues, our main outcomes consist of age-specific measures but we also separately consider criminal charges at a given point in the educational cycle.

We consider two types of age-specific crime measures: one outcome measures whether an individual has been charged with a 'crime *at* a given age' from age 15 and onwards. This is a memoryless measure, which simply informs about the tendency to commit crime at any given age. It is particularly useful for detecting sudden changes in the crime-age profile caused by school starting age. Our other type of outcome measures whether an individual has been charged with a 'crime *at or before* a given age' and in this way keeps track of earlier incidences. This is convenient if one wants to address more permanent effects on crime. Because of considerable recidivism,[19] both measures are required to give a full picture of the consequences of school starting age on the crime-age curve. We might see negative effects of school starting age on 'crime at a given age' but not at 'crime at or before a given age' if those committing the crime are simply the same individuals. Conversely, we could see effects on 'crime at or before a given age' and not on 'crime at a given age' if school starting age has a longer-lasting effect on criminal behavior. It is clearly important to be able to distinguish between these scenarios.

Due to space considerations we will sometimes focus on crime at or before a given age only but the full set of descriptives and formal results is available on request. In addition to our main analyses, sub-analyses show results for types of crime (property crime and violent crime) and number of crimes to address differential effects on the intensive and extensive margin.

Figure 5 illustrates means of our main outcome variables. The figure replicates the well-known age pattern where criminal activity peaks at ages 19-20 (Gottfredson and Hirschi (1990)). For girls, two percent are charged with a crime at age 19, while for boys 11 percent are charged with a crime at

[19] Recidivism is between 20% and 67% for boys and between 20% and 90% for girls, depending on age.

age 19 after which age the fraction declines. All over the age range, the proportion charged with a crime is higher for individuals who are old-for-grade compared to individuals who are young-for-grade. Our empirical analysis will reveal to what extent this reflects selection.

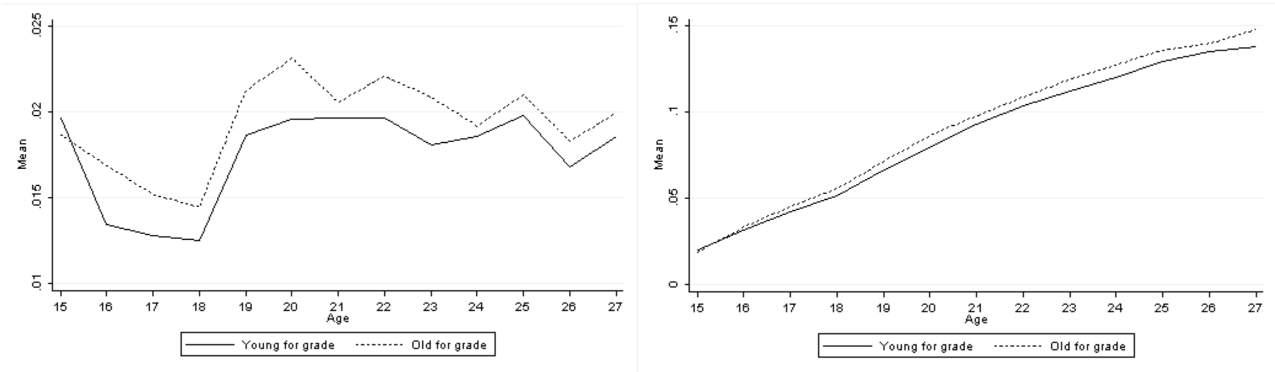
FIGURE 5

MAIN OUTCOME VARIABLES: CRIME ACROSS AGES

ANY CRIME AT GIVEN AGE

ANY CRIME BEFORE OR AT AGE

GIRLS



BOYS

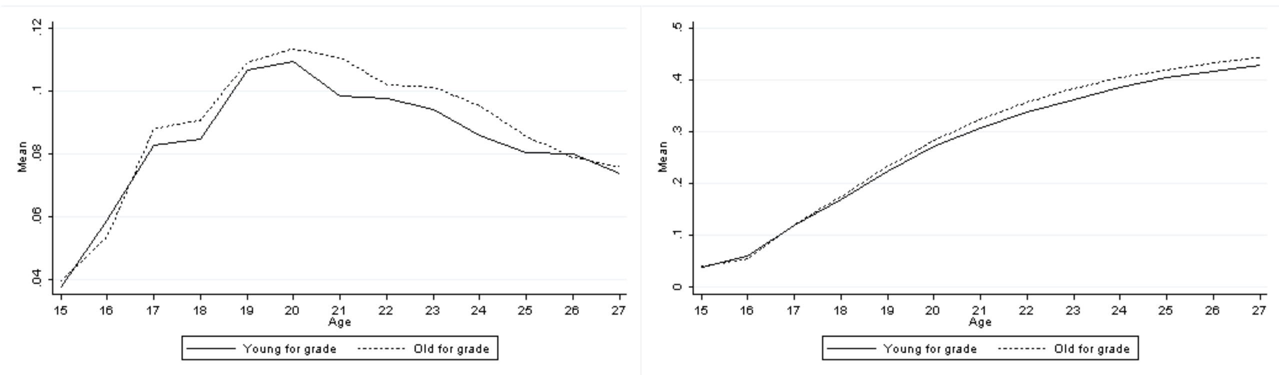


TABLE 1

MEANS OF SELECTED OUTCOME VARIABLES BY TYPES OF CRIME

Criminal charge (0/1)	Girls					Boys				
at or before age:	Any	Property	Violence	Other	Nobs	Any	Property	Violence	Other	Nobs
15	0.019	0.018	0.002	0.001	48,546	0.039	0.033	0.006	0.007	50,383
16	0.032	0.029	0.003	0.003	48,546	0.081	0.059	0.014	0.029	50,383
17	0.044	0.037	0.005	0.006	48,546	0.140	0.081	0.025	0.077	50,383
18	0.054	0.043	0.007	0.010	48,546	0.189	0.101	0.036	0.119	50,383
19	0.069	0.049	0.009	0.021	43,668	0.243	0.118	0.047	0.174	45,368
20	0.083	0.053	0.010	0.034	39,037	0.290	0.131	0.057	0.222	40,606
21	0.096	0.055	0.010	0.045	34,559	0.327	0.141	0.065	0.262	36,012
22	0.106	0.056	0.011	0.057	30,209	0.358	0.147	0.070	0.295	31,405
23	0.116	0.056	0.011	0.068	26,093	0.383	0.151	0.074	0.323	26,937
24	0.124	0.057	0.012	0.075	22,125	0.402	0.154	0.076	0.345	22,781
25	0.133	0.057	0.012	0.086	18,240	0.417	0.156	0.079	0.362	18,723
26	0.138	0.057	0.012	0.092	14,630	0.429	0.156	0.081	0.375	14,949
27	0.143	0.057	0.011	0.100	11,045	0.439	0.157	0.082	0.388	11,273

To give a better sense of the nature of the crime committed, Table 1 summarizes the distribution of crime at or before a given age across three types of crime: property crime, violent crime, and other crime.[20] Throughout the age distribution, boys are three times more likely to have been charged with a crime than girls. At the youngest ages, property crimes tend to be most prevalent, but after age 18 when the individuals in the sample gradually acquire a driver's license, other crimes including traffic incidents accumulate. For girls, other crime dominates from age 22 onwards, while for boys it dominates already from age 18.

Violent crimes comprise the most severe crimes. The most common examples are ordinary assaults, aggravated assaults, threats, and violence towards public servants. 80 percent of convictions for violence result in imprisonment or probation for boys and 67 percent for girls. Property crimes and other crime are typically less severe crime. The most frequent examples of property crime are shoplifting, burglary, and vandalism. A quarter of all convictions for property crime result in imprisonment or probation for boys, while this number falls below ten percent for girls. The

[20] Due to space considerations, we have chosen only three broad categories of crimes. Different classifications would be possible.

category of other crimes is dominated by traffic related crime (50% for girls and 90% for boys) such as driving a car without a license) while the second largest category is drug or weapon related crime (e.g. selling drugs or possession of illegal weapons). Convictions for other crimes rarely lead to imprisonment.

Table A3 shows mean crime outcomes by birth-month and gender. Those born in December tend to be more likely to have been charged with a crime compared to those born in January. When we consider whether an individual has been charged with a crime at a given age, we see that boys born in December are more likely to have been charged with a crime at each age from 15 to 22, while the pattern is more scattered for girls (top panel). This outcome will be important for our analysis of incapacitation. When we consider the accumulated measure, namely whether or not the individual has been charged with a crime at or before a given age, we see that the difference is significant up until age 24 (mid panel). As argued above, the accumulated outcome is more informative about potential catching up effects and other long run effects.

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. All specifications include the distance in days to the cut-off. Remaining estimates may be found in Table A4 in the Appendix.

In line with Figure 3, 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.[21] Table A5 in the appendix shows the 1st stage results by different subpopulations. The table shows that cutoff is highly significant for all subpopulations in question.

B. Crime results: 2SLS

Figure 6 shows our main estimation results. The left hand side figures show the estimation results for crime at a given age and the right hand side figures show the estimation results for crime at or before a given age. We find that being old-for-grade leads to a significant reduction in the propensity to commit crime at each age until age 19 for boys but only at age 15 for girls. Estimates at older ages are primarily negative for girls but become very close to zero for boys. Note that individuals who are young-for-grade turn 15 during their final year in comprehensive school, while individuals who are old-for-grade turn 16. Individuals who are young-for-grade turn 18 or 19, while individuals who are old-for-grade turn 19 or 20 during their final year in high school (depending on whether they took the optional 10th grade or not). This pattern in our results is supportive of the incapacitation hypothesis. It suggests that compulsory school is protective against crime for girls, while also high school is protective against crime for boys.

When we instead look at the propensity to commit crime *at or before* a given age, we find a statistically significant effect for girls until age 19. After age 19, estimates are again primarily negative but imprecise. Hence it seems that for girls, a higher school starting age initially reduces crime and we see no catching up at older ages. Estimates for boys are significant until age 22, after

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

which they become very close to zero. For boys therefore, we see a longer-lasting initial effect that eventually fade. The fading effects suggest that crime at the extensive margin is aligned to key life events rather than age. If criminal behavior instead was fixed to age, any effects of school starting age to crime should shift the crime-age profile in vertical direction. However, Figure 6 show that the effects fade in the long run when old-for-graders' and young-for-graders' educational attainment and life-course converge, which is consistent with a parallel shift to the crime-age profile. Moreover, as the effects only go to zero and not above suggest that the crime-age profile is shifted in both vertical and horizontal direction.

The 'delay' in crime is large relative to the mean. The share of girls with any criminal charges at or before 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.Appendix A, Table A6 shows detailed estimation results for crime at or before a given age where we gradually add control variables. In Table A7 in the appendix we distinguish between types of crime. For girls the significant effects of school starting age on crime at or before a given age are mainly driven by the impact on violent crimes, while for boys the effects are primarily driven by the impact on property crimes, although the effects on the two other categories of crime are significant for some ages. Appendix A, Figure A3 show the same pattern when we consider years of completed schooling as an outcome: initially we find large reductions for boys and girls, which fade and go to zero as age increase.

TABLE 2

FIRST STAGE ESTIMATION RESULTS

CHILDREN BORN IN DECEMBER AND JANUARY

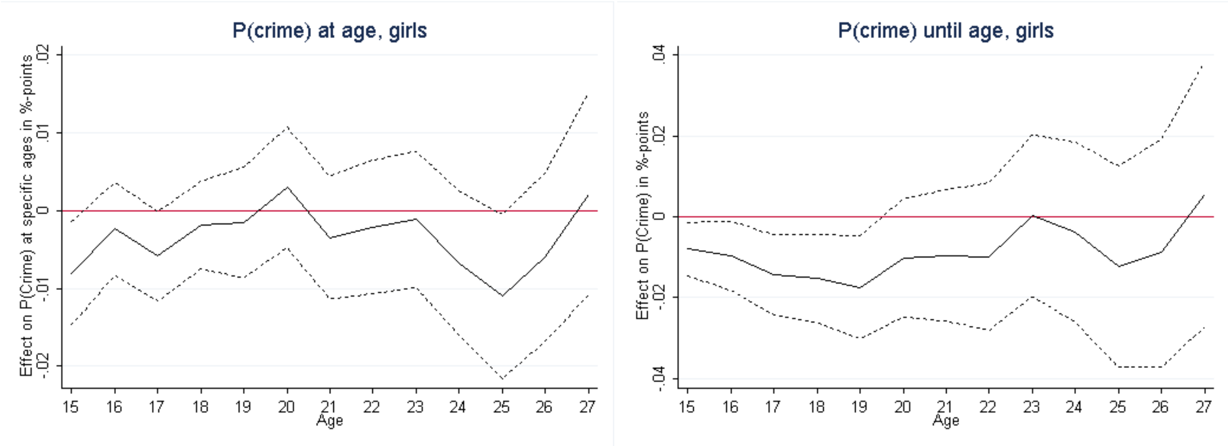
Variables	Girls		Boys	
	OLS	OLS	OLS	OLS
Born in Dec	-0.34 ***	-0.34 ***	-0.21 ***	-0.21 ***
	-24.99	-25.23	-17.46	-17.59
Born in Dec 1981	-0.03	-0.03	-0.15 ***	-0.15 ***
	-1.67	-1.61	-8.47	-8.54
Born in Dec 1982	-0.08 ***	-0.08 ***	-0.15 ***	-0.15 ***
	-3.72	-3.83	-8.47	-8.49
Born in Dec 1983	-0.06 **	-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 FE, distance to cut-off	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.

FIGURE 6

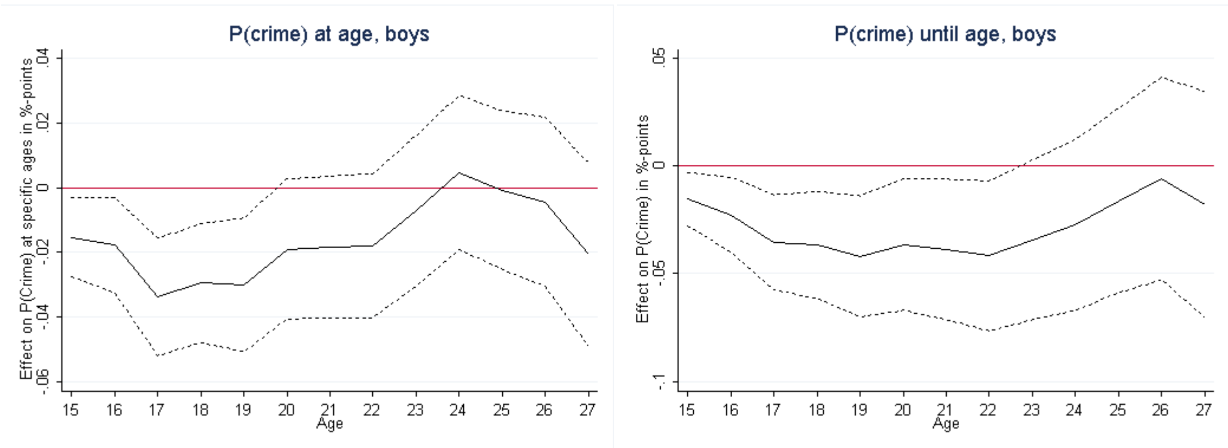
ESTIMATION RESULTS: CRIME ACROSS AGE

GIRLS



Note: The dashed lines indicate 95% confidence intervals.

BOYS

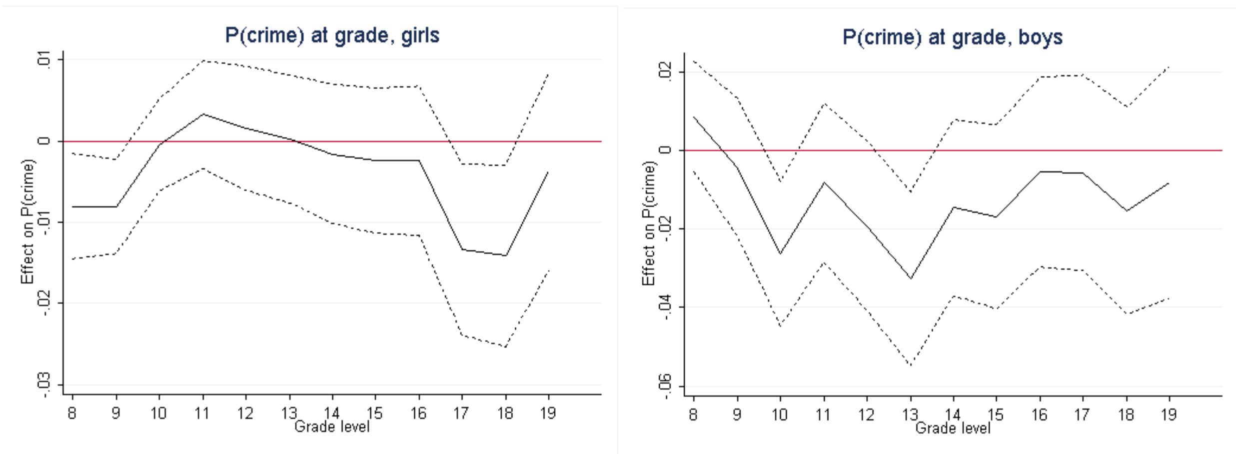


Note: The dashed lines indicate 95% confidence intervals.

Just as Black et al. (2011) show in the case of test scores, the way we align crime is extremely important for our conclusions. If one instead considers criminal charges at a given point in the educational cycle, one could potentially conclude that school starting age only has minor impacts on crime outcomes. Figure 7 shows results that align children in terms of grades instead of age. If the effects of school starting age only originated from delayed life-course, aligning grade should nullify the effects. The figure shows that age-gradient is smaller than for the age-aligned crime; it nullifies the effect at some but not all grade levels. For girls, the effect of school starting age on crime is significantly negative for the two final years in comprehensive school (grade levels 8 and 9). This corresponds to the significantly negative results at age 15 in our main analysis above. For boys the effect is only significant at the transitions from one level to the next in the educational cycle (grade levels 10 and 13). These results speak in favor of our interpretation of the results presented in Figure 6; being old-for-grade actually lowers crime in itself and not only in connection with changes to life events.

FIGURE 7

ESTIMATION RESULTS: CRIME ACROSS GRADES

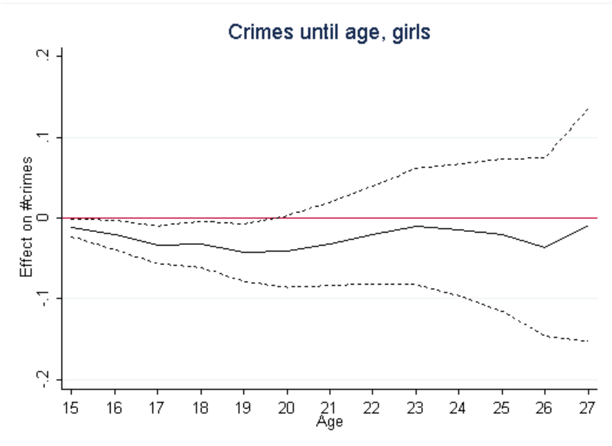


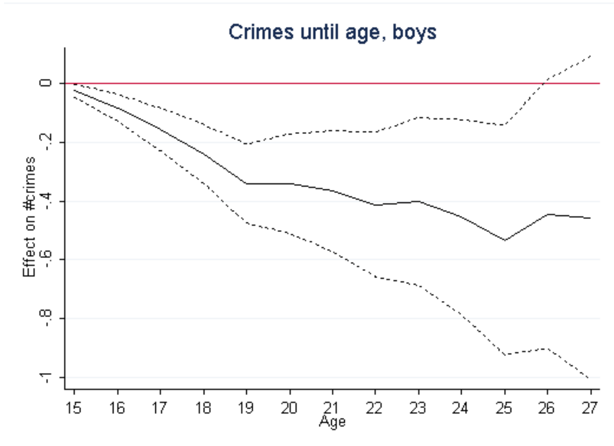
Note: The dashed lines indicate 95% confidence intervals.

In Figure 8, we consider the effects on crime at the intensive margin. The figure presents the effects of increasing school starting age on the number of charges at or before a given age. The estimates for girls are significant in the same age range as the results for the indicator variable above. This is likely because the majority of girls only commit very few crimes. For boys, however, this exercise reveals interesting additional insights that were not clear from the simple indicator analysis: estimates are much larger and effects last long into the twenties. In their mid-twenties young men who started school later as a consequence of being born in January have been charged with half a crime less than those who did not. This is a substantial effect which has large consequences for both to the offenders and potential victims. Moreover, these large and persistent effects also show that the intensive margin crime-age profile for boys is much more related to age than the extensive margin crime-age profile we investigated earlier. Criminal behavior is thus not only determined by either age or key events but by both in interaction; it matters at what age one is exposed to different key events.

FIGURE 8

ESTIMATION RESULTS: NUMBER OF CRIMES AT OR BEFORE AGE X





Note: The dashed lines indicate 95% confidence intervals.

C. Generalization of results using a cutoff as IV

The estimates we have presented so far are Local Average Treatment Effects; i.e. average effects for the group of compliers that shift school starting age due to the cutoff date. However, the effects of being old-for-grade may be heterogeneous across different margins of treatment. The presence of significant heterogeneity could imply that some experience large effects of being old-for-grade while others do not, implying that our estimated LATEs cannot be generalized to the entire population and in turn reduce the external validity of our results.

Put in terms of our econometric framework from section III the potential outcome for treatment status j is $Y_j = M_j(X) + U_j,\ j = 0{,}1$ where $M_1(X) - M_0(X)$ are observed gains from being old-for-grade and $U_1 - U_0$ are unobserved gains from being old-for-grade. Heterogeneous effects imply that the unobserved gains $U_1 - U_0$ depends on the treatment probability $p(Z,X)$. In consequence, the estimated effects could differ from those presented earlier in the paper if we based the analysis on an IV that affect a different margin on *p(Z,X)* (Brinch, Mogstad & Wiswall (2014)). We can investigate this formally by testing whether:

(7) $$E[Y|D=1,Z=1] - E[Y|D=0,Z=1] = E[Y|D=1,Z=0] - E[Y|D=0,Z=0]$$

$$=> \ E[U_1 - U_0|\ Z=1] = E[U_1 - U_0|\ Z=0]$$

for each outcome. If we reject the hypothesis, the effect of being old-for-grade is heterogeneous in our sample. Table A8 in the appendix reports the results of the test for external validity of the LATEs estimates using the cut-off date unconditionally and conditional on year of birth.[22] The table shows that we reject the naught of homogeneous effects in our sample for most outcomes. To elucidate these heterogeneous effects we estimate Marginal Treatment Effects of being old-for grade.

[22] E(D|Z=1) and E(D|Z=0) differ across years which allow us to test for external validity of the LATE at several margins of treatment.

Complier Characteristics

In Table 3 below we summarize the average characteristics of the compliers (those who shift to being old-for-grade because of the IV, see Abadie (2003)) along with the average characteristics of the full sample to serve as comparison. Families who change the school start decision as a consequence of being born in January rather than December tend to have more favorable characteristics: parents are more often living together, birth weight is higher, and parents have higher education and stronger attachment to the labor market. Thus, at the margins at which we have support, low socio-economic families with low ability children are less likely to be affected by the annual cutoff we use for causal inference (cf. areas of common support in Figures A4 and A5).

TABLE 3

COMPLIER CHARACTERISTICS

Variable	Girls			Boys		
	Sample mean	Complier mean	t-stat	Sample mean	Complier mean	t-stat
Immigrant	0.04	0.02	4.99 ***	0.04	0.02	3.91 ***
Parents married/cohabiting	0.79	0.81	-2.72 **	0.79	0.83	-3.45 ***
Apgar score=9	0.18	0.17	1.39	0.19	0.17	1.25
Apgar score=8	0.07	0.06	2.74 **	0.07	0.08	-2.39 *
Apgar score lower	0.08	0.09	-0.12	0.10	0.08	1.78
Birth weight, grams	3349.43	3414.11	-5.90 ***	3473.43	3589.40	-7.21 ***
Gestational length, weeks	39.55	39.61	-1.59	39.47	39.62	-2.67 **
Mother:						
- Months of schooling	137.41	139.08	-2.49 *	137.75	142.92	-5.87 ***
- Completed HS or equvalent	0.29	0.29	0.08	0.30	0.34	-3.51 ***
- Unemployed	0.13	0.11	3.00 **	0.13	0.10	2.31 *
- Out of the labor force	0.11	0.10	1.29	0.11	0.09	1.59
- Age at birth of first child	24.85	24.91	-0.76	24.92	25.18	-2.18 *
Father:						
- Months of schooling	140.15	143.26	-3.85 ***	140.38	146.86	-5.71 ***
- Completed HS or equvalent	0.19	0.18	2.58 **	0.20	0.23	-3.29 ***
- Unemployed	0.08	0.07	1.67	0.07	0.06	2.28 *
- Out of the labor force	0.06	0.04	4.51 ***	0.06	0.04	3.26 **

Note: Standard errors are bootstrapped, * indicates $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

Marginal treatment effects

This section presents the estimation results and confidence intervals for the marginal treatment effects of being old-for-grade. One critical issue is the discrete nature of our exogenous variation: Our instrument is binary or at best discrete with 22 different points in $P(Z)$ when we condition on our 11 different cohorts. In order to accomodate this issue, we use the framework introduced in Brinch, Mogstad, and Wiswall (2014) to estimate MTE using discrete instruments. We estimate $P(Z,X)$ using the full sample and subsequently estimate:

(8) $$E\left[Y_j|P(Z,X)=p, X=x, D=j\right]=M_j(X)+K_j(p,x),\ j=0,1$$

The MTE is then:

(9) $$MTE=M_1(X)-M_0(X)+k_1(p,s)-k_0(p,x),$$

where $k_1=p\cdot\frac{dK_1}{dp}$ and $k_0=-(1-p)\cdot\frac{dK_0}{dp}$. We can identify $K_j(p,X)$ up to an 11th order polynomial of $P(Z,X)$ by assuming separability between M_j and U_j in the outcome equation.[23]

The following figures show the estimation results and confidence intervals for the marginal treatment effects of increased school starting age. In all figures, the x-axis shows the normalized term U_D (see section III) which, in the settings of a generalized Roy model, may be interpreted as the net-utility of being old-for-grade given the outcome and information-set parents use to decide their child's school starting age. Both the outcome of interest in the choice equation and the information set is unobserved from the researcher's perspective. We will discuss what might enter the choice equation in section III.

[23] We choose order of K_j and bandwidth by a K-fold cross-validation with 10 folds. In practice we only find few differences between polynomials above second order. Larger order polynomials affect the curvature of the lowest and highest ranges of support – however, nowhere near significantly as the standard errors in these ranges are very large.

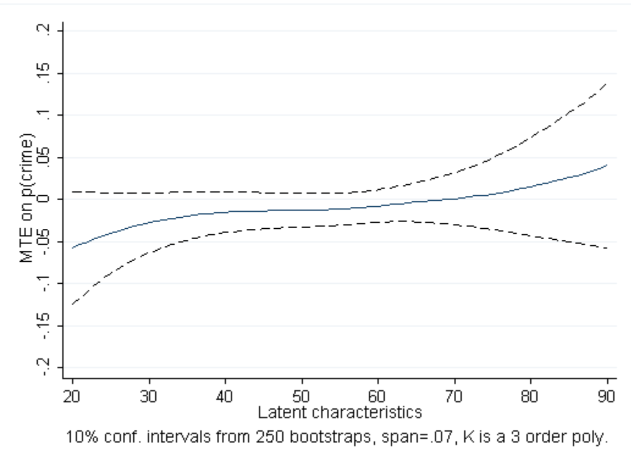
Of course, we can only estimate the MTE within the intervals of common support. Figures A4 and A5 show these areas for girls (.20-.90) and boys (.40-.95). The lack of common support for low values of latent characteristics for boys arises because the least able children born in December always delay school start. The weighted average of the different MTEs across the entire set of latent characteristics equals the 2SLS estimates from Table A6.

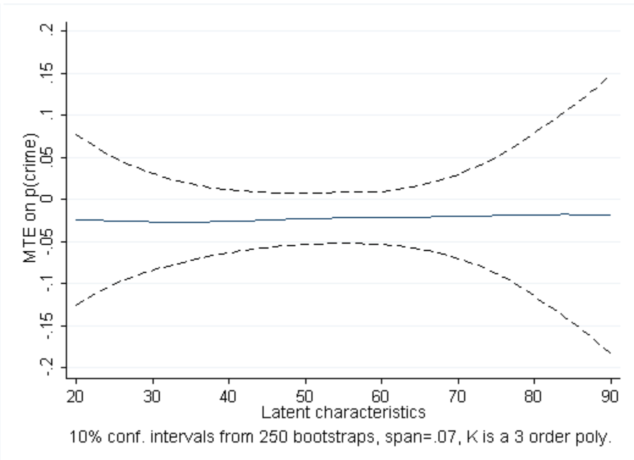
FIGURE 9

MARGINAL TREATMENT EFFECT: CRIME AT OR BEFORE AGE 15 AND 18

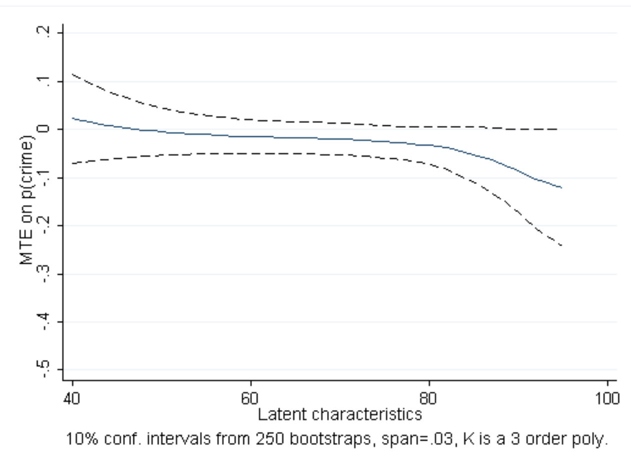
AGE 15 | AGE 18

GIRLS





BOYS



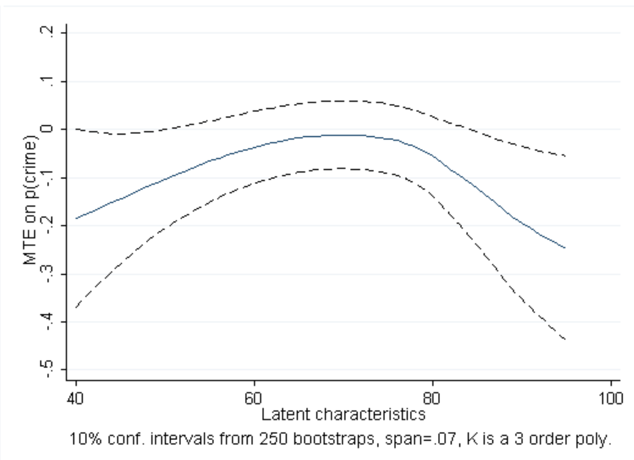


Figure 9 shows the estimated MTEs of being old-for-grade on two selected outcome variables: crime at or before age 15 (left) and age 18 (right) for girls (top) and boys (bottom), respectively. The upper panels of Figure 9 show no significant slopes of the MTEs on crime before or at age 15 and 18, respectively. The bottom panels of Figure 9 shows some heterogeneity for boys. For crime before or at age 15 (bottom left panel), the MTE is downward sloping. Even though the slope is not significant it suggests that those with high levels of latent characteristics benefit the most in terms of reducing crime. When we consider crime before or at age 18 (bottom right panel), the MTE is inversely U-shaped, and the MTE is only significantly different from zero at high latent abilities. The slope in the region where U_D is 70-95 is significant, implying essential heterogeneity. Those with high U_D also benefit the most in terms of reducing crime between age 16 to 17. The MTE predict a reduction to the probability of committing crime by up to 25 percentage points for those with highest U_D.

D. Potential Mechanisms and Effects on Alternative Outcomes.

This section first attempts to shed light on some of the different channels through which school starting age may affect crime outcomes. Specifically, we further investigate the importance of incapacitation and also consider the role played by the relative age of peers as in Black et al. (2013). We next address effects on an alternative range of outcomes. Specifically, we discussed above that parents may choose to enroll their child later in school even if there are no long run effects on income, for example. Because school starting age is linked to grades, it may also be linked to the quality of and consumption value associated with the type of degree. Finally, municipality based variation in culture may impact on parents' choices. This section investigates these hypotheses.

We first address incapacitation. In Figure A6, we investigate how school starting age affects crime across the week. For boys, the effect is driven by crime committed during weekdays. The effect is statistically significant for most weekdays until age 19 or 20, while it is not to the same extent significant for Saturdays and Sundays. For girls, the effect during the weekdays is not statistically different from zero. One exception is Fridays, where the effect is borderline significant for 15-17 year-olds. We interpret these findings as supportive of incapacitation effect for boys throughout high school, which the age pattern of the main results already pointed at above. For girls the effects are smaller and the mechanism is more subtle: the effect is driven by violent crimes, crimes taking place on Fridays and the effect dies out around completion of compulsory school. Thus is appears that girls who are old-for-grade are less likely to be involved in this type of crime while they are still in school.

Table A9 analyzes the effect of the age of peers in line with Fredriksson and Öckert (2005) and Black et al. (2013). 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*.[24] 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 variable.[25] This is in line with the findings in the mentioned previous studies for Norway and Sweden.

We finally investigate other mechanisms that may explain why parents choose to postpone school start even when effects on children's primary long-term outcomes are moderate in size.

Table 4 shows the estimated effects of being old-for-grade on alternative outcomes such as grades and type of degree that may enhance the consumption value of school for children and parents. We

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

[25] Compare the results presented in Table A6 to Table A9.

find that the impact of school starting age on standardized math grades is statistically significant and large in magnitude.[26] This is in line with previous studies supporting large age-at-test effects (Crawford, Dearden and Meghir (2010) and Black, Devereux and Salvanes (2011)). If such grades make a difference for educational preferences or feasible choices, they may influence long term outcomes. Indeed we do see that girls are more likely to enroll in one of the selective and competitive Medical Schools, while boys tend to obtain a slightly longer education if they are old-for-grade as a consequence of being born on the other side of the cut-off., Organization (or effort) grades are not affected for boys or girls.

To investigate the variation in the enforcement of the stipulated school starting age across municipalities we look at the distribution of predicted school starting age using a rich set of observable characteristics against the actual school starting age across municipalities. We find little relationship between the predicted and actual school starting pattern on municipal level. Moreover, in ten percent of all municipalities, less than 68% (49%) of all boys (girls) born +/-30 days from the cut-off are old-for-grade, while in another ten percent of all municipalities more than 84% (67%) are old-for-grade conditional of observables. This suggests that local school start culture and legal enforcement of the regulations may play a role for the parents' decision.

[26] Danish grades are not affected (not shown). Figure A7 shows the corresponding MTEs. We find that the MTE for math and organization grades are constant for boys, while for girls we see a significant downward slope for effort grades (or organization). The girls with the lowest U_D experience around two standard deviations higher grades from being old-for-grade whereas the girls with the highest U_D are unaffected.

TABLE 4

EFFECTS OF SCHOOL STARTING AGE ON OTHER OUTCOMES

Variable	Girls			Boys		
	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Grades						
- Math	-0.101 ***	0.288 ***	0.283 ***	-0.070 **	0.302 ***	0.275 ***
	0.012	0.036	0.033	0.015	0.051	0.047
- Effort	-0.143 ***	0.014	0.012	-0.106 ***	0.030	0.017
	0.011	0.032	0.032	0.015	0.053	0.051
# Observations	27,909	27,909	27,909	27,974	27,974	27,974
Years of schooling (by age 27)	-0.653 ***	-0.044	-0.042	-0.426 ***	0.234 **	0.165
	0.054	0.135	0.124	0.056	0.147	0.135
# Observations	11,045	11,045	11,045	11,273	11,273	11,273
College enrollment						
- Med School	-0.011 ***	0.024 *	0.025 *	-0.008 ***	0.010	0.011
	0.001	0.011	0.011	0.004	0.010	0.010
- Law School	-0.010 ***	-0.012	-0.011	-0.004 ***	-0.015	-0.014
	0.002	0.124	0.012	0.001	0.011	0.011
# Observations	30,209	30,209	30,209	31,405	31,405	31,405
Controls						
- Yearly cut-off FE, distance to cut-off	X	X	X	X	X	X
- Covariates			X			X

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

VI. Conclusion

This paper uses Danish register-based data to investigate the effect of school starting age on crime-age profiles 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 until age 20. In addition, boys experience a persistent reduction in the *number* of crimes committed. We show that crime at the extensive margin is largely driven by life events whereas crime at the intensive margin is a complex function of both age and life-course.

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 20 years, and this effect stems from crime (primarily property crime) taking place during the weekdays. Girls who are old-for-grade are less likely to be charged until they turn 18, and this effect stems from crime (mainly violent crimes) taking place on Fridays. For boys we find significant effects of school starting age on the accumulated number of crimes at or before a given age throughout the twenties. For girls the effects on accumulated crime measures die out which suggests that school starting age only influences the criminal debut.

For boys mainly property crime is reduced while for girls violent crime is reduced. Importantly, we find evidence of essential heterogeneity for boys but not for girls. 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.

Our results suggest that increasing school starting age could lower crime – more so for boys than for girls. Yet, our findings do not necessarily 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

Abadie, A., (2003): Semiparametric instrumental variable estimation of treatment response models, *Journal of Econometrics*, 113(2), 231-263

Anderson, M. (forthcoming): In School and Out of Trouble? The Minimum Dropout Age and Juvenile Crime. Forthcoming in the *Review of Economics and Statistics.*

Bedard, K. and E. Dhuey (2006): The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects. *Quarterly Journal of Economics* 121 (4): 1437-1472.

Björklund, A. and 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.

Brinch, C., M. Mogstad and M. Wiswall (2014): Beyond LATE with a Discrete Instrument: Heterogeneity in the Quantity-Quality Interaction of Children. Manuscript.

Cook, P. J. and S. Kang (2014): Birthdays, schooling, and crime: New evidence on the dropout-crime nexus. Manuscript presented at ASSA 2014.

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. and D. Lubotsky (2009): Kindergarten Entrance Age and Children's Achievement Impacts of State Policies, Family Background, and Peers. *Journal of Human Resources* 44(3), 641-683.

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., M. S. Morrill, 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.

Gottfredson, M. R., and T. Hirschi (1990): *A General Theory of Crime*. Stanford, CA: Stanford UP.

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

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

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.

Sampson, R. J. and J. H. Laub (1995): *Crime in the Making: Pathways and Turning Points through Life*. Harvard UP.

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

MEANS OF OUTCOME VARIABLES ACROSS CUT-OFF

Any criminal charge (0/1)	Girls				Boys			
at age:	December	January	Difference	Nobs	December	January	Difference	Nobs
15	0.021	0.017	**0.003**	48,546	0.042	0.036	**0.007**	50,383
16	0.016	0.015	0.001	48,546	0.058	0.051	**0.007**	50,383
17	0.015	0.013	**0.002**	48,546	0.093	0.081	**0.012**	50,383
18	0.014	0.013	0.001	48,546	0.095	0.084	**0.010**	50,383
19	0.021	0.020	0.001	43,668	0.114	0.102	**0.012**	45,368
20	0.021	0.022	-0.001	39,037	0.117	0.108	**0.009**	40,606
21	0.021	0.019	0.002	34,559	0.112	0.103	**0.008**	36,012
22	0.022	0.020	0.001	30,209	0.105	0.097	**0.008**	31,405
23	0.020	0.019	0.001	26,093	0.101	0.097	0.005	26,937
24	0.020	0.018	0.003	22,125	0.093	0.092	0.002	22,781
25	0.023	0.018	**0.005**	18,240	0.085	0.082	0.003	18,723
26	0.019	0.016	0.002	14,630	0.081	0.077	0.004	14,949
27	0.019	0.020	-0.001	11,045	0.080	0.071	0.009	11,273

Any criminal charge (0/1)	Girls				Boys			
at or before age:	December	January	Difference	Nobs	December	January	Difference	Nobs
15	0.021	0.017	**0.003**	48,546	0.042	0.036	**0.007**	50,383
16	0.035	0.030	**0.004**	48,546	0.086	0.076	**0.010**	50,383
17	0.047	0.041	**0.006**	48,546	0.147	0.133	**0.014**	50,383
18	0.057	0.051	**0.006**	48,546	0.196	0.181	**0.015**	50,383
19	0.073	0.065	**0.008**	43,668	0.252	0.234	**0.018**	45,368
20	0.086	0.081	0.005	39,037	0.298	0.282	**0.017**	40,606
21	0.098	0.093	0.005	34,559	0.336	0.318	**0.018**	36,012
22	0.109	0.104	0.005	30,209	0.368	0.348	**0.020**	31,405
23	0.116	0.115	0.001	26,093	0.391	0.374	**0.018**	26,937
24	0.125	0.123	0.003	22,125	0.410	0.395	**0.016**	22,781
25	0.136	0.130	0.006	18,240	0.423	0.410	0.013	18,723
26	0.140	0.136	0.004	14,630	0.433	0.424	0.009	14,949
27	0.142	0.144	-0.001	11,045	0.446	0.433	0.013	11,273

Note: Bold indicate that the difference is statistically significant at the 5%-significance level.

TABLE A4

FIRST STAGE RESULTS (SUPPL. TO TABLE 2 IN MAIN TEXT)

Variables	Girls		Boys	
Immigrant	-0.07	***	-0.10	***
	-5.62		-9.00	
Apgar score=9	0.01		0.00	
	1.01		-0.73	
Apgar score=8	0.01		0.00	
	1.47		-0.34	
Apgar score<8	0.01	+	0.00	
	1.70		0.63	
Birthweight (kg)	-0.06	***	-0.04	***
	-13.80		-11.97	
Gestational length	-0.07	***	-0.03	*
	-3.55		-1.98	
Gestational length sq (/100)	0.08	**	0.04	+
	3.25		1.94	
Mother's months of schooling (/100)	0.07	***	0.06	
	3.40		3.29	
Mother's months of schooling sq (/10000)	-0.05	***	-0.04	***
	-5.75		-4.42	
Father's months of schooling (/100)	0.04	*	0.04	**
	2.49		2.83	
Father's months of schooling sq (/10000)	-0.02	**	-0.02	*
	-3.11		-2.45	
Mother has completed high school	-0.02	***	0.00	
	-3.72		-0.41	
Father has completed high school	-0.06	***	-0.05	***
	-8.85		-8.27	
Mother was unemployed	0.02	**	0.00	
	2.65		0.86	
Mother was out of labor force	0.00		-0.03	***
	-0.45		-4.70	
Father was unemployed	0.02	***	0.00	
	2.61		0.52	
Father was out of labor force	0.00		0.00	
	-0.26		-0.04	
Mother's age at first child(/100)	0.04		0.19	***
	0.77		3.71	
Parrents were married	0.00		0.01	
	0.82		1.23	

TABLE A4 (CONTINUED)

Variables	Girls	Boys
New year 1981-1982	-0.13 ***	-0.08 ***
	-9.28	-6.16
New year 1982-1983	-0.09 ***	-0.06 ***
	-6.57	-4.83
New year 1983-1984	-0.11 ***	-0.07 ***
	-7.39	-5.71
New year 1984-1985	-0.12 ***	-0.07 ***
	-8.72	-5.54
New year 1985-1986	-0.09 ***	-0.05 ***
	-6.65	-4.42
New year 1986-1987	-0.07 ***	-0.04 **
	-5.16	-3.20
New year 1987-1988	-0.03 *	-0.02
	-2.32	-1.64
New year 1988-1989	-0.03 *	0.00
	-2.43	0.21
New year 1989-1990	-0.02	0.00
	-1.15	0.26
New year 1991-1992	0.01	0.01
	0.77	0.97
New year 1992-1993	0.00	0.01
	-0.14	0.46
Dec distance to cutoff	0.00 ***	0.00 ***
	-14.97	-17.35
Jan distance to cutoff	0.00 ***	0.00 ***
	9.23	4.03
Observations	48,546	50,383

TABLE A5

FIRST STAGE RESULTS FOR SUBPOPULATIONS

Sample selection	Girls December	Boys December
Mother's educ. shorter than High School	-0.35 ***	-0.25 ***
	64.65	51.85
Mother's educ. at least High School	-0.38 ***	-0.27 ***
	61.83	51.80
Mother in labor force	-0.37 ***	-0.27 ***
	85.88	70.74
Mother out of labor force	-0.31 ***	-0.23 ***
	24.25	19.58
Mothers age at first child < 25	-0.36 ***	-0.25 ***
	58.88	48.55
Mothers age at first child >= 25	-0.37 ***	-0.27 ***
	66.97	54.75
Parrents are married	-0.32 ***	-0.23 ***
	35.23	28.84
Parrents are not married	-0.38 ***	-0.27 ***
	82.14	67.58
Not immigrant or descendant	-0.37 ***	-0.26 ***
	88.22	72.18
Immigrant or descendant	-0.28 ***	-0.26 ***
	12.86	12.03
Below median birth weight	-0.39 ***	-0.29 ***
	64.04	58.86
At or above median birth weight	-0.34 ***	-0.23 ***
	62.25	44.25
Controls		
- Yearly cut-off FE, distance to cut-off	X	X
- Covariates	X	X

Note: For simplicity the table shows results for only 1 cutoff instrument and not year-by-year instruments.
* indicates p < 0.05, ** p < 0.01, and *** p < 0.001.

TABLE A6

SELECTED DETAILED ESTIMATION RESULTS: CRIME AT OR BEFORE AGE X

Outcome	Girls				Boys			
Any criminal charge (0/1) at or before age:	OLS	2SLS	2SLS	Nobs	OLS	2SLS	2SLS	Nobs
15	-0.004 ***	-0.009 *	-0.008 *	48,546	-0.006 **	-0.021 ***	-0.015 *	50,383
	0.001	0.003	0.003		0.002	0.006	0.006	
16	-0.001	-0.010 *	-0.010 *	48,546	-0.010 ***	-0.033 ***	-0.023 *	50,383
	0.002	0.004	0.004		0.003	0.009	0.009	
17	0.000	-0.015 **	-0.014 **	48,546	-0.005	-0.049 ***	-0.036 **	50,383
	0.002	0.005	0.005		0.004	0.012	0.011	
18	0.001	-0.016 **	-0.015 **	48,546	0.000	-0.052 ***	-0.037 **	50,383
	0.002	0.006	0.006		0.004	0.013	0.013	
19	0.002	-0.019 **	-0.018 **	43,668	0.004	-0.062 ***	-0.042 **	45,368
	0.002	0.007	0.007		0.005	0.015	0.014	
20	0.004	-0.011	-0.010	39,037	0.006	-0.057 ***	-0.036 *	40,606
	0.003	0.008	0.007		0.005	0.016	0.016	
21	0.002	-0.011	-0.010	34,559	0.010 +	-0.059 ***	-0,0387 *	36,012
	0.003	0.008	0.008		0.006	0.017	0.017	
22	0.003	-0.011	-0.010	30,209	0.012 *	-0.064 ***	-0.042 *	31,405
	0.004	0.009	0.009		0.006	0.019	0.018	
23	0.006	-0.001	0.000	26,093	0,0163 *	-0.055 **	-0.034 +	26,937
	0.004	0.010	0.010		0.007	0.020	0.019	
24	0.006	-0.005	-0.004	22,125	0.015 *	-0.046 *	-0.027	22,781
	0.004	0.012	0.011		0.007	0.021	0.020	
25	0.007	-0.014	-0.012	18,240	0.012	-0.036 +	-0.016	18,723
	0.005	0.013	0.013		0.008	0.023	0.022	
26	0.005	-0.010	-0.009	14,630	0.015 +	-0.025	-0.006	14,949
	0.006	0.015	0.014		0.009	0.025	0.024	
27	0.010	0.006	0.005	11,045	0.014	-0.034	-0.018	11,273
	0.007	0.017	0.017		0.010	0.028	0.027	
Controls								
- Yearly cut-off FE, distance to cut-off	X	X	X		X	X	X	
- Covariates			X				X	

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

TABLE A7

ESTIMATION RESULTS: TYPES OF CRIME

Outcome	Girls				Boys			
Any criminal charge (0/1) at or before age:	Violence	Property	Other	Nobs	Violence	Property	Other	Nobs
15	-0.001	-0.006 +	-0.001	48,546	0.002	-0.013 *	-0.006 *	50,383
	0.001	0.003	0.001		0.002	0.006	0.003	
16	-0.003 +	-0.005	-0.003 *	48,546	-0.001	-0.022 **	-0.018 **	50,383
	0.001	0.004	0.001		0.004	0.008	0.006	
17	-0.003 +	-0.009 +	-0.005 *	48,546	-0.009 +	-0.033 ***	-0.027 **	50,383
	0.002	0.005	0.002		0.005	0.009	0.009	
18	-0.006 **	-0.007	-0.005 +	48,546	-0.013 *	-0.035 ***	-0.031 **	50,383
	0.002	0.005	0.003		0.006	0.010	0.011	
19	-0.008 ***	-0.010 +	-0.003	43,668	-0.016 *	-0.043 ***	-0.032 *	45,368
	0.002	0.006	0.004		0.007	0.011	0.013	
20	-0.008 **	-0.011 +	0.003	39,037	-0.017 *	-0.037 **	-0.032 *	40,606
	0.003	0.006	0.005		0.008	0.012	0.014	
21	-0.008 **	-0.009	0.001	34,559	-0.010	-0.039 **	-0.034 *	36,012
	0.003	0.006	0.006		0.009	0.012	0.016	
22	-0.006 +	-0.009	-0.002	30,209	-0.012	-0.046 ***	-0.036 *	31,405
	0.003	0.007	0.007		0.010	0.013	0.017	
23	-0.004	-0.005	0.003	26,093	-0.010	-0.045 **	-0.029	26,937
	0.003	0.007	0.008		0.010	0.014	0.018	
24	-0.004	0.000	-0.005	22,125	-0.008	-0.049 **	-0.025	22,781
	0.004	0.008	0.009		0.011	0.015	0.020	
25	-0.004	-0.002	-0.013	18,240	-0.009	-0.052 **	-0.019	18,723
	0.004	0.009	0.011		0.012	0.016	0.021	
26	-0.007	-0.006	-0.003	14,630	-0.010	-0.047 **	-0.001	14,949
	0.005	0.010	0.012		0.013	0.018	0.024	
27	-0.004	0.002	0.000	11,045	0.005	-0.047 *	-0.017	11,273
	0.005	0.011	0.014		0.015	0.020	0.026	
Controls								
- *Yearly cut-off FE, distance to cut-off*	X	X	X		X	X	X	
- *Covariates*			X				X	

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

TABLE A8

TEST FOR EXTERNALVALIDITY OF LATE

	Girls Criminal charges at age				Boys Criminal charges at age			
	15	16	17	18	15	16	17	18
All years								
1982	***	*			***			
1983	***		+	*	***			
1984	***				***			
1985	***							
1986								
1987		*			+	**	*	+
1988								
1989								
1990								
1991							+	*
1992		+	+	+				+
1993						*		

Note: + indicates $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

TABLE A9

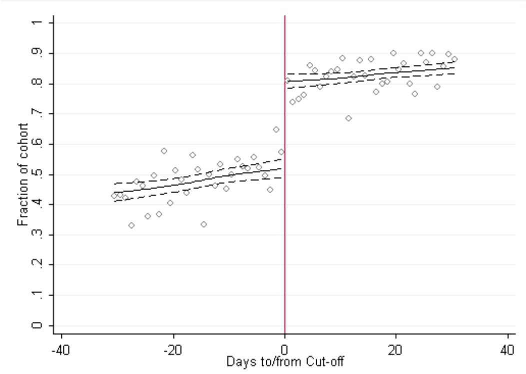
ABSOLUTE AND RELATIVE EFFECTS OF SCHOOL STARTING AGE ON CRIME

	Girls			Boys		
Outcome Criminal charges at or before age:	Old-for-grade	Peer age	Nobs	Old-for-grade	Peer age	Nobs
15	-0.008 *	0.020	48,546	-0.016 *	0.011	50,383
	0.003	0.028		0.006	0.042	
16	-0.010 *	-0.020	48,546	-0.023 *	0.015	50,383
	0.004	0.036		0.009	0.060	
17	-0.014 **	-0.003	48,546	-0.035 **	-0.031	50,383
	0.005	0.042		0.011	0.076	
18	-0.015 **	0.003	48,546	-0.038 **	0.067	50,383
	0.006	0.046		0.013	0.085	
19	-0.017 **	-0.045	43,668	-0.044 **	0.146	45,368
	0.007	0.054		0.014	0.097	
20	-0.010	-0.044	39,037	-0.038 *	0.084	40,606
	0.007	0.061		0.016	0.108	
21	-0.009	-0.084	34,559	-0.040 *	0.103	36,012
	0.008	0.068		0.017	0.115	
22	-0.009	-0.083	30,209	-0.043 *	0.044	31,405
	0.009	0.076		0.018	0.123	
23	0.001	-0.081	26,093	-0.035 +	0.037	26,937
	0.010	0.083		0.019	0.130	
24	-0.001	-0.132	22,125	-0.027	0.003	22,781
	0.012	0.094		0.021	0.138	
25	-0.008	-0.226	18,240	-0.014	-0.072	18,723
	0.013	0.102		0.022	0.150	
26	-0.004	-0.186	14,630	0.005	-0.336 *	14,949
	0.015	0.113		0.025	0.166	
27	0.009	-0.119	11,045	-0.006	-0.323 +	11,273
	0.017	0.130		0.028	0.186	
Controls						
- Yearly cut-off FE, distance to cut-off	X			X		
- Covariates	X			X		

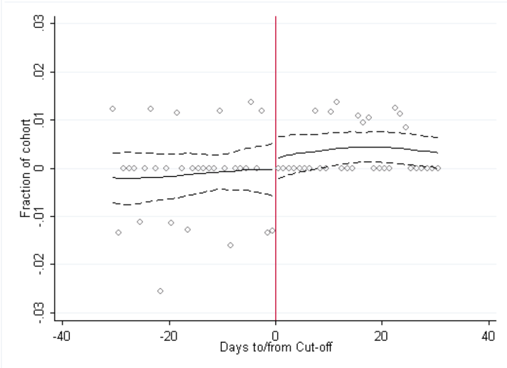
FIGURE A1

CHANGES IN FRACTION OLD-FOR-GRADERS IN 2007 AROUND THE CUTOFF, GIRLS

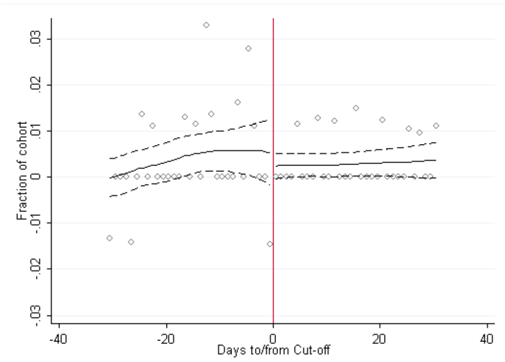
BEFORE SCHOOL START



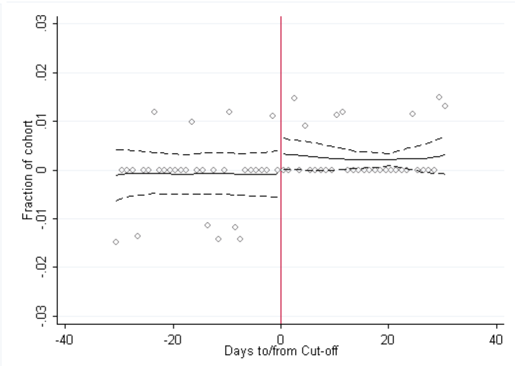
GRADE 1



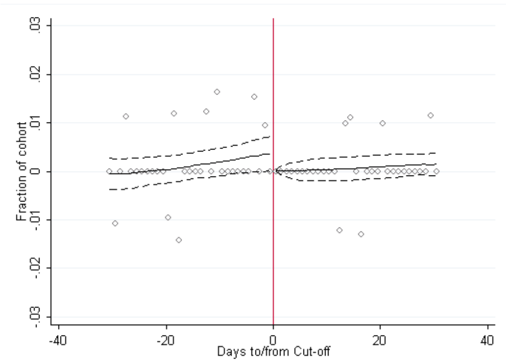
GRADE 2



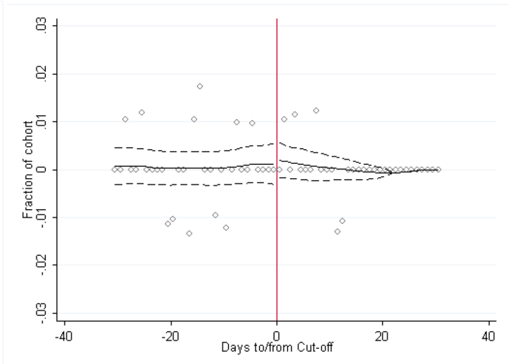
GRADE 3



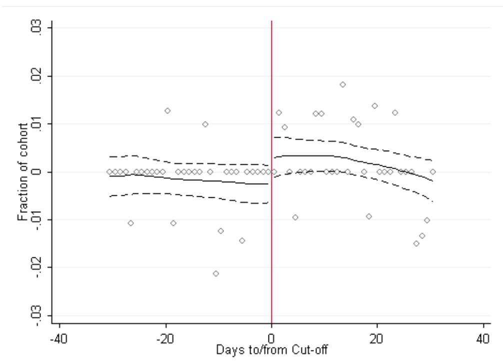
GRADE 4



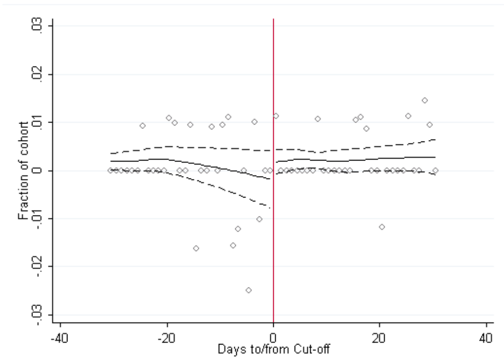
GRADE 5



GRADE 6



GRADE 7

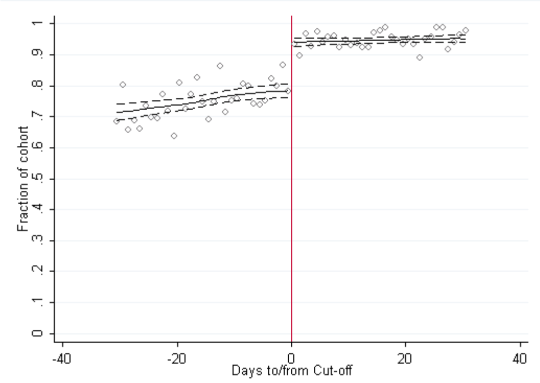


Note: The dashed lines indicate 95% confidence intervals.

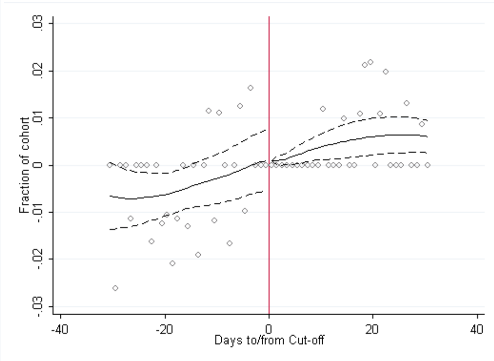
FIGURE A2

CHANGES IN FRACTION OLD-FOR-GRADERS IN 2007 AROUND THE CUTOFF,BOYS

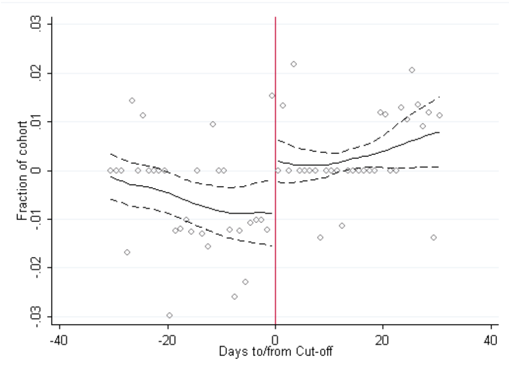
BEFORE SCHOOL START



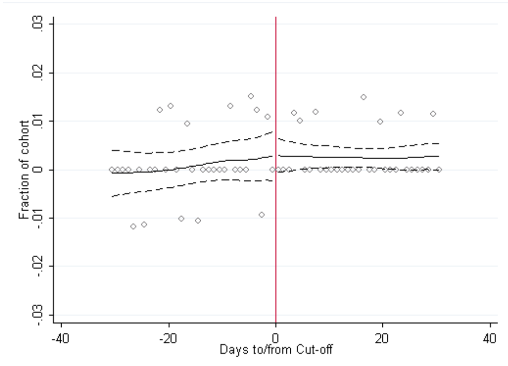
GRADE 1



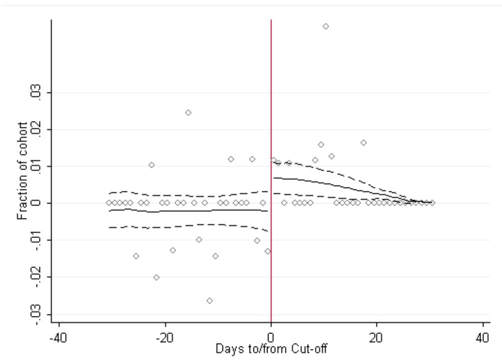
GRADE 2



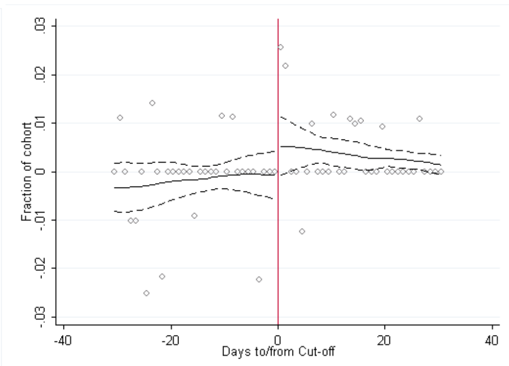
GRADE 3



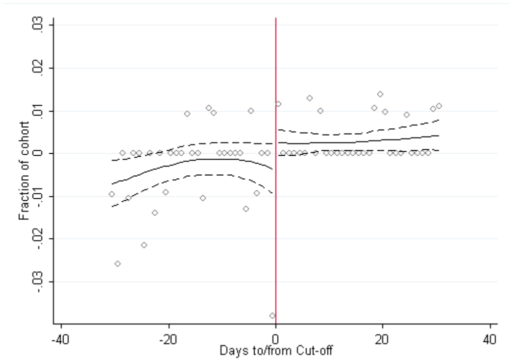
GRADE 4



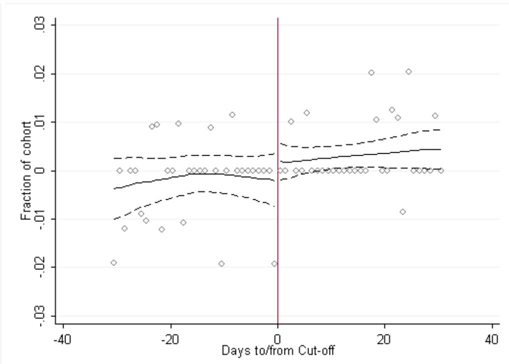
GRADE 5



GRADE 6



GRADE 7



Note: The dashed lines indicate 95% confidence intervals.

FIGURE A3

EFFECT OF BEING OLD-FOR-GRADE ON YEARS OF COMPLETED SCHOOLING

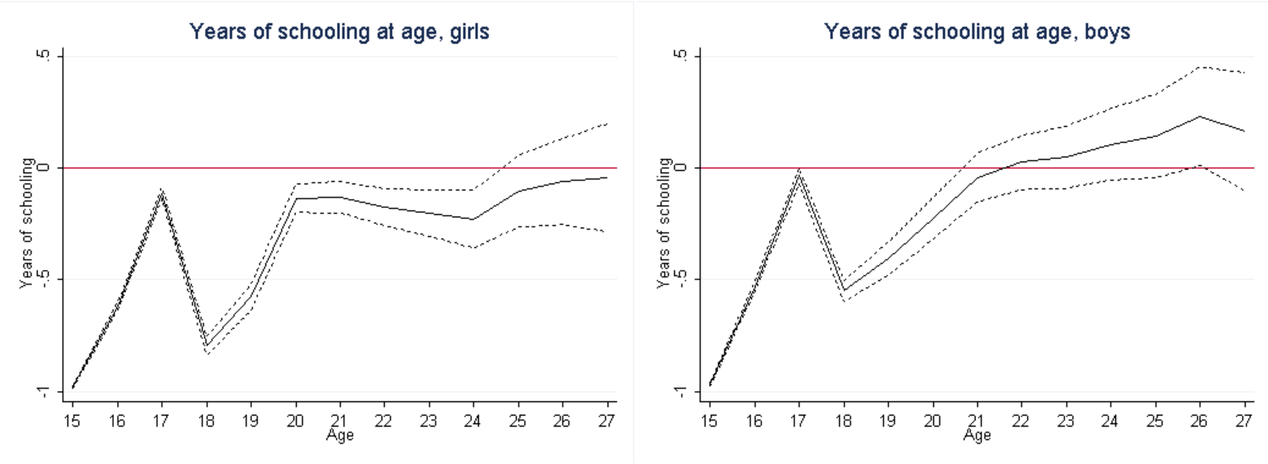


FIGURE A4

AREA OF COMMON SUPPORT, GIRLS

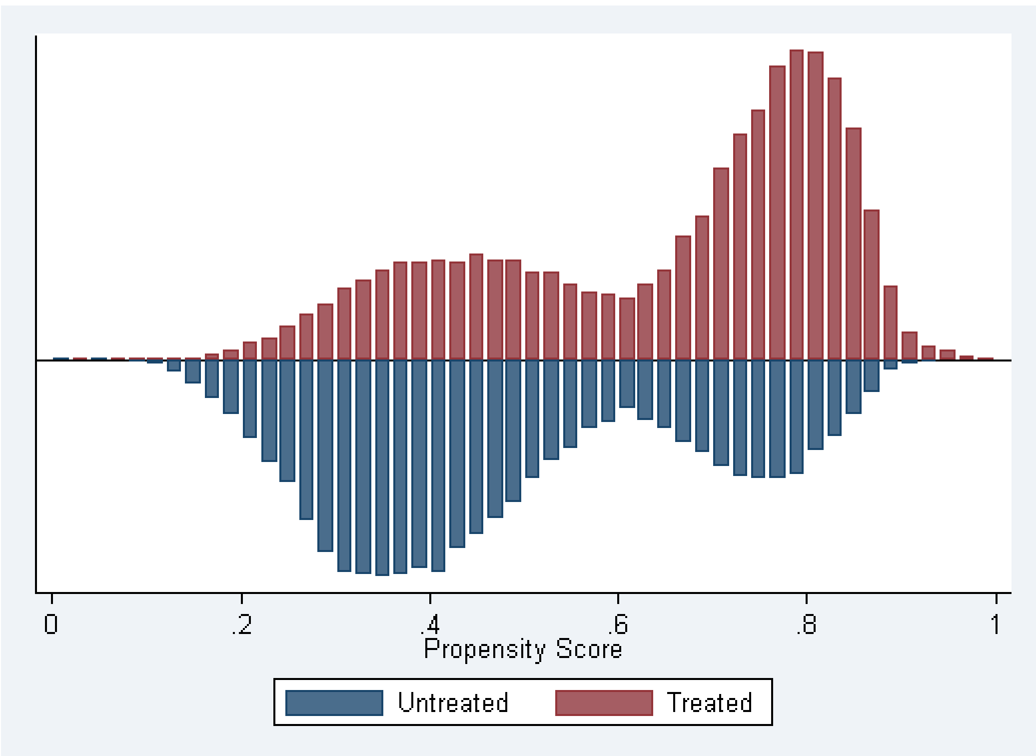


FIGURE A5

AREA OF COMMON SUPPORT, BOYS

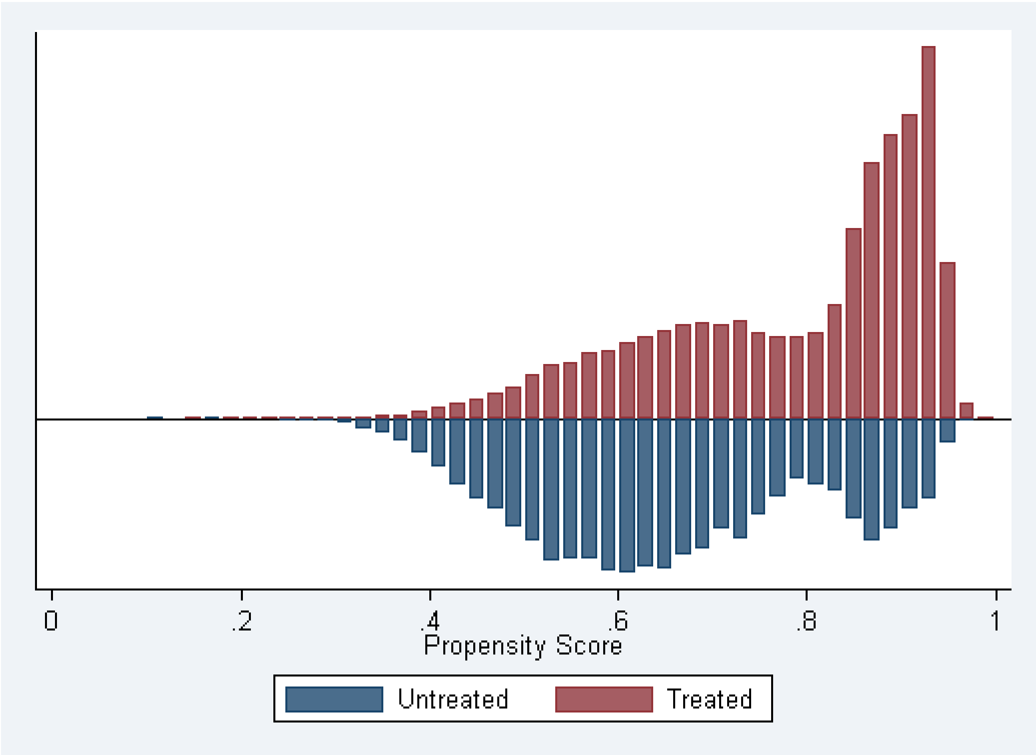
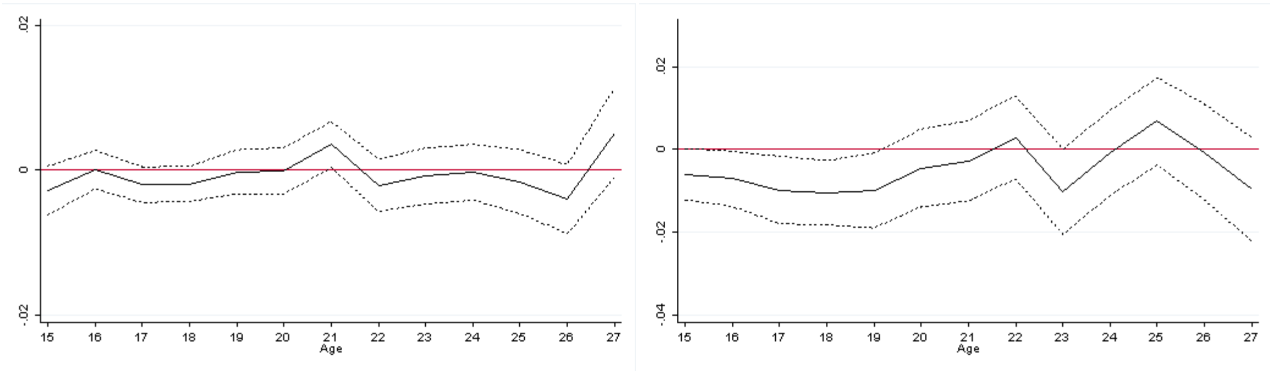


FIGURE A6

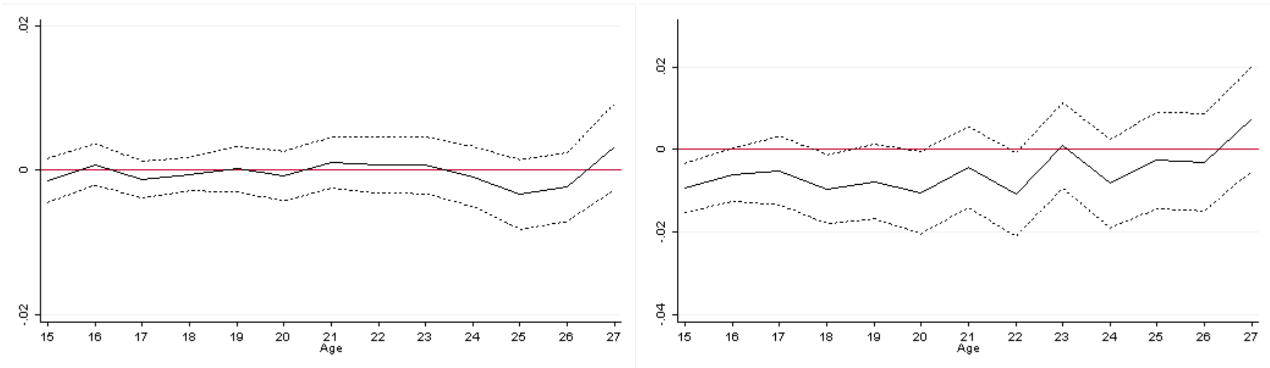
ESTIMATION RESULTS: CRIME ACROSS THE WEEK

GIRLS BOYS

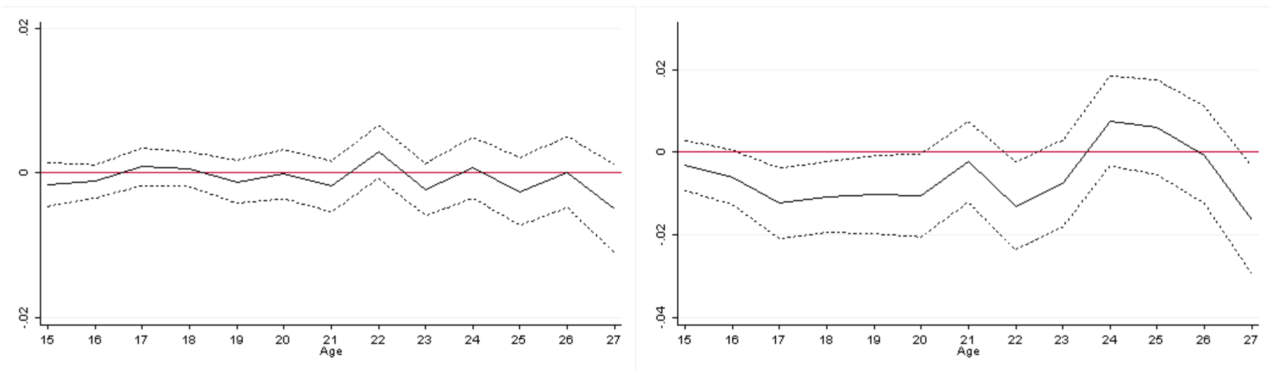
MONDAY



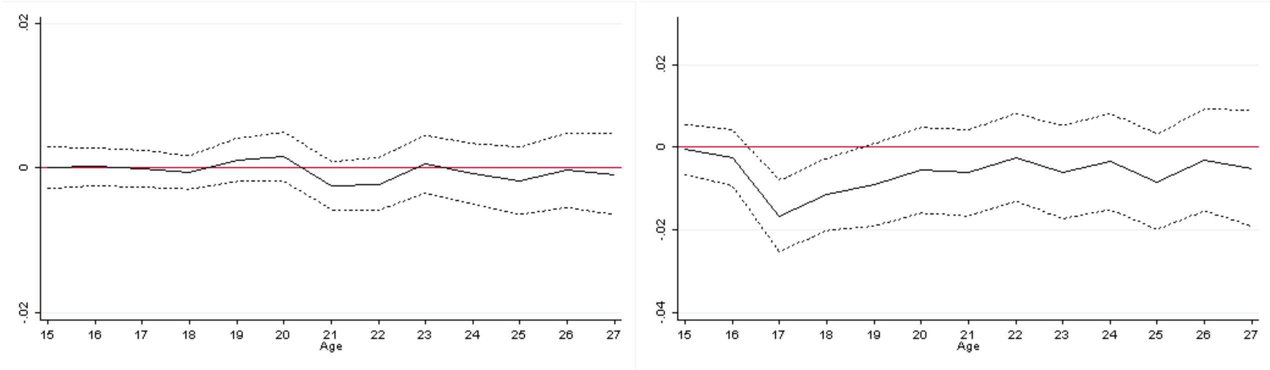
TUESDAY



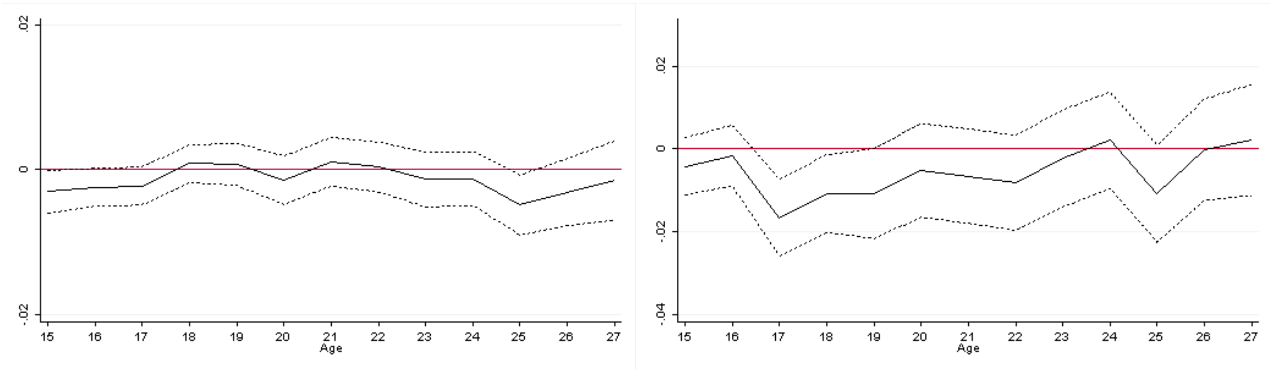
WEDNESDAY



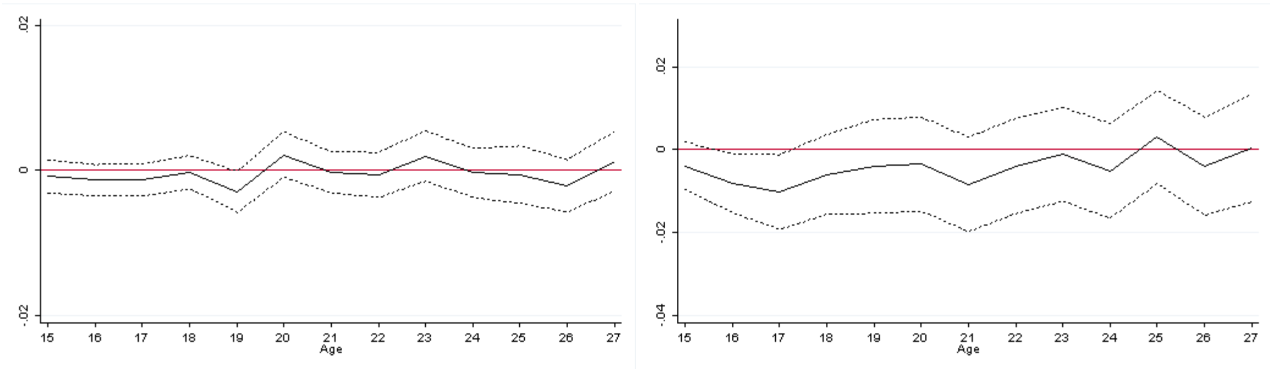
THURSDAY



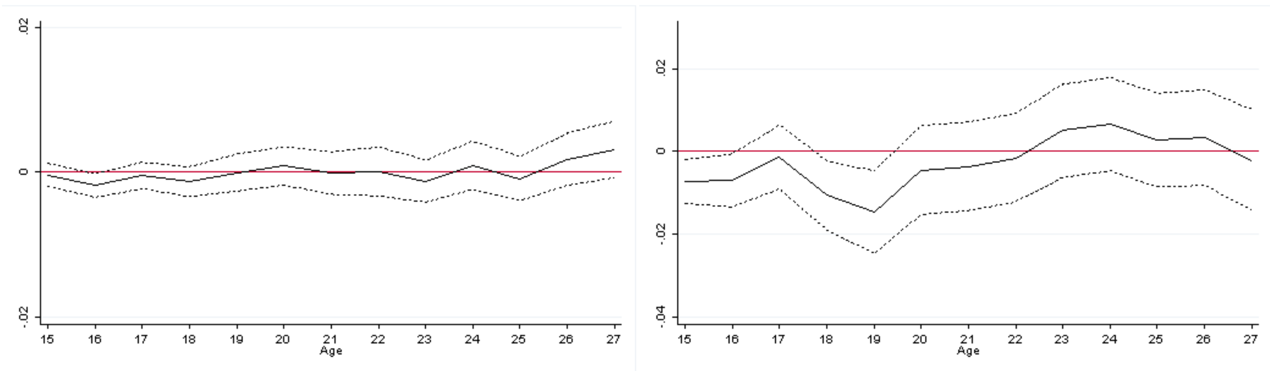
FRIDAY



SATURDAY



SUNDAY



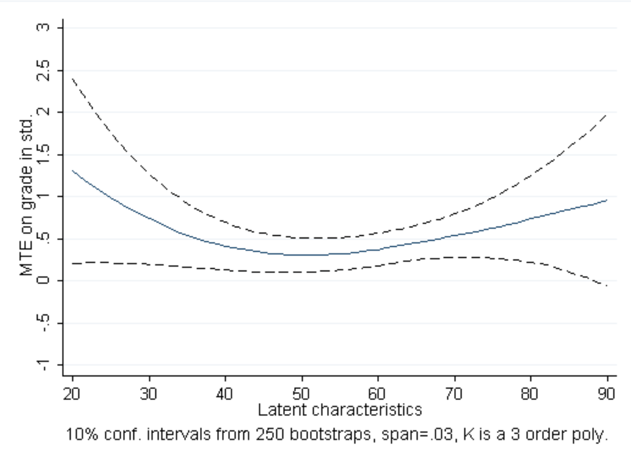
Note: The dashed lines indicate 95% confidence intervals.

FIGURE A7

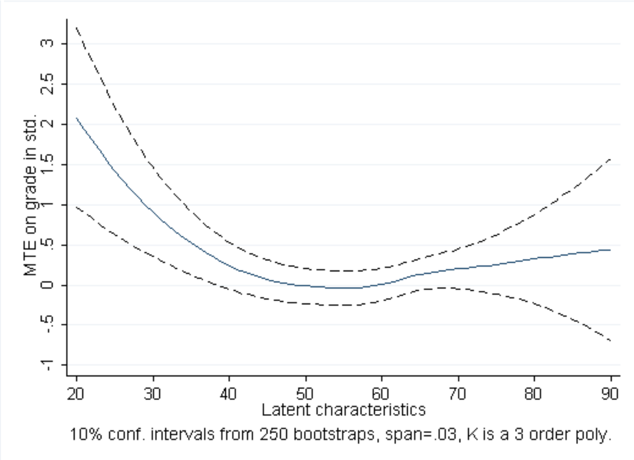
MARGINAL TREATMENT EFFCTS: OTHER OUTCOMES

GIRLS

Math grades (standardized)

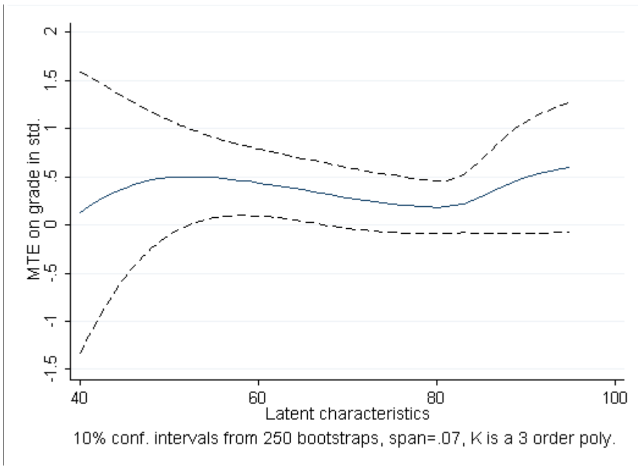


Organization grades (standardized)



BOYS

Math grades (standardized)



Organization grades (standardized)