Education and Youth Crime: Exploring the Causal Relationship

Olivier Marie[*], Tanja Traag[**], and Rolf van der Velden[*]

This Version: September 2009

Very preliminary, please do not quote without permission of the authors.

Abstract

In this paper we explore the causal relationship between education and youth crime. We carefully consider the endogeneity problems of schooling and offending decision making which plague this relationship. We suggest solving these issues by instrumenting educational attainment by the difference in the chances of obtaining a secondary qualification brought about by timing of birth of students. We use a unique database which matches a large survey of Dutch youths to administrative data on arrest for these individuals. Using different specification of discrete choice models, we find a relatively small 5 percent negative link between obtaining a starter's qualification on the probability of arrest aged 16 to 19. The first stage of our IV approach suggests that relatively older students are more likely to leave secondary school without a valid degree. The reduced form yields much larger impacts of schooling on crime than we previously estimated. A secondary school degree reduces the chance of arrest by at least 12 percentage points. We believe this to be a more accurate measure of the school crime causal relationship and suggest it should be considered to develop education policies to incentivise students to obtain formal qualification.

[*]Research Centre for Education and the Labour Market (ROA), Maastricht University, The Netherlands
[**]Central Bureau for Statistics (CBS), Heerlen, The Netherlands

1 - Introduction

There is a large body of evidence that the individuals who leave school early or without any formal qualification have much higher probability of being criminally active and/or arrested (Freeman 1996, 1999; and Lochner 2004). Low wages and high unemployment have been shown to increase crime (Grogger 1998; Raphael and Winter-Ebmer 2001; Gould et al 2002; and Machin and Meghir 2004) and, since education improves future labour market opportunities, this can partly explain the negative relationship observed between schooling and subsequent adult crime. This evidence suggests that research should focus on explaining contemporaneous crime and education decision making among youths. Still there is very little work which has investigated this relationship in order to show if this is a causal relationship or a simple correlation due to unobserved characteristics. The reason is mainly that certain individual traits, such as patience and attitude to risk, may influence *both* the decision to invest in education *and* participate in crime. It is therefore very difficult to establish if obtaining a qualification would change the offending behaviour of youths who would otherwise have not chosen to stay at school because of these unobserved characteristics. Even with high quality data on educational attainment and criminal activity of youths it is necessary to methodologically consider the importance of this endogeneity problem to investigate the school-crime relationship.

In this paper we attempt to address these difficulties by considering a possible instrumental variable (IV) which induces differences in schooling but is uncorrelated to other factors which directly affect criminal behaviour (e.g. preferences or ability). Intuitively, this approach exploits differences in educational attainment across individuals that arise in response to factors that have no direct effect on criminal decisions. An ideal instrument would randomly assign some youth to drop out and others to finish secondary school with some qualification. Then, comparing the differences in crime rates across these groups would identify the causal effect of secondary education completion on crime. Here we consider if individuals born earlier or later in the year are likely to do better or worse at school than their classmate. This relationship was shown to be prevalent in a large number of previous international studies (Angrist and Krueger 1991, 1992; Bedard and Dhuey 2006; Crawford et al 2007; and Grenet 2009). We argue that the random timing of birth of individuals over the year is a good IV since it is unlikely to influence the criminal behaviour of individuals except through the influence it has on their educational attainment. Differences in relative age should therefore enable us to identify the impact of schooling on criminal participation. We believe that even if it does not yield extremely precise estimates it should least indicate the magnitude of the causal relationship.

To do so we use detailed information from the 1999 VOCL survey of 18,000 randomly selected Dutch students entering secondary school. It contains a large number of observed characteristics of the pupils and their parents which could influence youth criminal behaviour. This data was matched to yearly police arrest data for these individuals up to age 19-20 and also to administrative data on student academic achievement at that age. This enables us to distinguish between the youths who have left school with and without formal qualification and those who have been suspected of a crime or not. We first observe that individuals who did not drop out of secondary school are more than four times less likely to be arrested by the police between 16 and 19 years old. These two groups of youths are statistically different on almost all their observable characteristics. We gradually control for this individual information in probit models. We find that the impact of obtaining some qualification on the probability of an arrest decreases from 13 to five percent once we include all the controls. This result is robust to including school level fixed effects and generating propensity scores to match individual on their observables. Because of the endogeneity problem we do not believe that this five percent reduction in arrest probability is the accurate measure of the impact of obtaining an educational qualification on criminal participation.

We therefore turn to our IV strategy and first illustrate the randomness of the distribution of the month of birth in the sample of individuals surveyed and in the Dutch population of a similar age. There is nevertheless a very marked difference in the distribution in the births of the students who have repeated a class in primary school. We can easily understand this phenomenon as resulting from the impact of timing of birth on the likelihood of grade retention early in life. This however means that we must carefully choose the measure of age we use as an instrument for educational attainment to account for this selection problem. We do so by calculating the assigned relative age (i.e. 12 - month of birth) as recommended by Bedard and Dhuey (2006). We find that students 11 months *older* than there peers in the first year of primary school have a seven to nine percent higher probability of leaving education without any qualification. These results are highly significant and correspond to the first stage of our IV analysis. Our findings from this exercise show that the impact of obtaining educational qualification on criminal behaviour is much higher than previously estimated. Leaving school with a diploma reduces your chances of arrest aged 16 to 19 by around 30 percentage points. We therefore conclude that the endogeneity of education and crime decision among youths leads to greatly under-estimating the benefits of not dropping out of secondary school on offending behaviour.

The rest of the paper is structured as follows. Section 2 discusses the school-crime relationship and our proposed estimation strategies. Section 3 describes the data and gives a number of descriptive characteristics. Section 4 reports and discusses the results. The final section concludes.

2 - The School-Crime Relationship and Estimation Strategies

A - The School-Crime Relationship

Following the human capital approach developed by Lochner (2004), it is simple to understand how schooling may influence subsequent adult crime participation. To the extent that education increases wage rates (and reduces the likelihood of unemployment), it increases the opportunity costs of crime and will tend to reduce post-school criminal activity. Higher wages raise the opportunity costs of crime in two distinct ways. First, since crime may require time to commit, that time cannot be used for other productive purposes like work. Here, it is useful to think of all of the time involved in planning a crime, locating a target and, potentially, evading detection and arrest. Second, each crime committed entails an expected period of incarceration, which is more costly for individuals with better labor market opportunities and wages. In the same spirit we can see how youth crime will tend to be decreasing in both contemporaneous and future wage rates. Higher contemporaneous wages increase the direct opportunity cost of committing crime, while higher future wages increase the costs associated with potential incarceration. Because education increases future wage rates, youth who are enrolled in school will be less likely to engage in crime than otherwise similar youth who are not in school.

The main problematic emerges because schooling is not exogenously determined. Youth will choose to enroll in school if they receive a net benefit from doing so; otherwise, they will not. Not only does an increase in returns to secondary and higher education reduce crime for all youth who would have attained these schooling levels in the first place, but it also causes more youth to finish high school and attend university, lowering their lifetime criminal activity as well. Since the benefits from schooling through higher lifetime earnings are delayed, youth who are more patient are more likely to attend school. More patient youth are also less likely to engage in crime, since the punishments tend to be delayed. Thus, differences in patience across the population will tend to lead to a negative relationship between education and crime. Population heterogeneity in preferences toward risk may also lead to a correlation between education and crime. If the rewards to school are risky as some economists suggest, more risk averse youth will

tend to quit school at earlier ages. Risk averse youth are also more likely to engage in crime regardless of their schooling, generating a negative correlation between crime and education.

These are some of the major issues which make it difficult to estimate the effects of schooling enrollment on contemporaneous crime. The fact that unobserved individual characteristics are partly responsible for the simultaneous decision youth make between the two activities make this a particularly daunting task which is often ill defined. One exception in recent research is the substantial contribution by Lochner and Moretti (2004) where the authors attempt to untangle the school-crime causal relationship. Their main methodological approach is to instrument enrollment decisions US states changes in compulsory schooling laws as in Angrist and Krueger (1991). With OLS they find that a one year increase in average education reduces arrest rates by 11 percent and high school graduation decreases probability of arrest by 70 percent. Their IV estimates yield relatively larger effects, although they are not statistically different. The methodology we will use in our analysis is relatively similar in the sense that it makes use of a tested instrument for school enrollment (i.e. timing of birth) to explore the causality of education attainment on criminal participation. We now describe our different identification strategies to estimate this effect.

B - Probit Models and PSM

The first strategy we implement is to obtain the simple estimates of the probability of arrest for individuals who have completed secondary education compared to those who have not. We can do this by obtaining probit estimates as in the following model:

$$\Pr(Arrest_i = 1) = \Phi(\beta Qual_i) \qquad (1)$$

where $\Phi(.)$ is the standard normal cumulative distribution function. *Arrest* is a dummy variable equal to 1 if an individual was arrested aged 16 to 19 and 0 otherwise, *Qual* is another dummy variable for this individual obtaining a secondary education qualification or not, and subscript *i* denotes each individual. This should give us a very rough estimate, β, of the correlation between obtaining a high school degree and being suspected of a crime by the police. To obtain a more relevant estimate of this coefficient by augmenting equation (1) as follows:

$$\Pr(Arrest_i = 1) = \Phi(\beta Qual_i + \sum_{k=1}^{K} \delta_k X_{ik}) \qquad (2)$$

This model includes a number, *K*, of controls for individual characteristics, *X*, which may also affect the probability of arrest of youths. The coefficients, δ_k, associated with each of these observed characteristics will be of interest to measure the relative impact they have on arrest

probability. Our main coefficient of interest will remain β which will now measure the partial correlation of obtaining a secondary education qualification on criminal participation, net of the effect of the individual controls we will have included. Finally we will augment this first probit estimation strategy by including school level fixed effects in the model. This will control for unobservable school characteristics which could have an influence on student offending behaviour and arrest probability.

One popular method to attempt to improve estimates in the statistical methods literature on treatment effects has been to resort to Propensity Score Matching (PSM). This methodology first suggested by Rosenbaum and Rubin (1983, 1984) allows for selection on observable characteristics in X to occur in a more flexible manner than in equation (2). The PSM method gives a score of the probability of being part of a treated group, here obtaining a secondary education degree, based on the following probit equation:

$$\Pr(Qual_i = 1) = \Phi(\sum_{k=1}^{K} \varphi_k X_{ik}) \qquad (3)$$

From (3) we generate propensity scores for each individual of the probability of completing high school. These scores can be used to match students with and without secondary school education with a similar score or their 'nearest neighbours' (i.e. who are as close as possible in terms of the X's). Once this is done we can again run a version of equation (2) to obtain an estimate of β but this time re-weighting each non-treated individual depending on how similar they are to their treated match depending on their propensity scores. We can expand this by also matching students across schools by including school level fixed effects in the model.

Although these methods should generate relatively precise estimates of β, these could still only measure a correlation and not the causal relationship of the impact of education attainment on crime we seek to observe. Because of the simultaneity of the decision making in school and crime behaviour in youth we have discussed above, we must consider an alternative methodology which takes this serious issue into account.

C- Instrumental Variable

To solve the problem of contemporaneous education and crime choices, we seek a factor which influences school enrolment but does not impact directly on offending behaviour or chances of arrest. One such instrument which has extensively been shown to impact on education outcomes is the timing of birth of students across the school year. In their seminal 1991 paper, Angrist and Krueger argue that in the US season of birth is related to educational attainment because

individuals born at the beginning of the year can drop out before those born at the end, simply because they reach the minimum leaving age earlier. They find that the students who were relatively older in a class cohort have on average significantly lower wages later in life and are able to estimate the labour market returns to schooling. As we have already discussed, current and future wage opportunities are likely to impact on criminal participation of youths. Consequently we argue that timing of birth of individuals will influence their offending behaviour of youths through the difference in educational attainment it generates.

Here we must note that relatively older students will reach an age at which individuals are more criminally active, the documented age-crime profile (Hirshi and Gottferdson 1983), earlier than their classmates. This problem can easily be dealt with by considering the arrest frequency of youth born at different times of the year at comparable ages. A more problematic issue is if being relatively younger or older in a class cohort influences criminal participation through other channels than educational attainment. One could assume that to be more mature physically or mentally could increase your chances of being more violent (e.g. a 'bully') or associate with peers likely to be more criminally active. We will for the moment assume that this link, not documented to our knowledge, have a minimal effect on offending behaviour in comparison to the impact on educational attainment of differences in relative age. Still this means we will be relatively conservative when interpreting our IV results until we are able to further investigate these issues.

Concretely our IV strategy will be a Two Stage Least Square (2SLS) estimation starting with the first stage equation for effect of quarter of birth, *QBirth*, on obtaining a secondary education qualification:

$$Qual_i = \alpha + \gamma QBirth_i + u_i \qquad (4)$$

where α is a constant and u_i is the error term. Coefficient γ will be a measure the impact of season on birth on schooling attainment. The reduced form relationship between arrest at age 16 to 19 and achieving a secondary education qualification is:

$$Arrest_i = \alpha + \beta QBirth_i + \varepsilon_i \qquad (5)$$

where ε_i is the error term and β measures the average impact of quarter of birth on arrest far all individuals *i*. This coefficient should give us an estimate of the causal relationship of education on criminal participation which takes into account the endogeneity issues our previous methodologies could not solve. Still we will seek to improve our IV modelling strategy considering two potential difficulties with the month of birth instrument.

First the crucial assumption that the timing of birth of children is random has recently been shown to be sometimes violated. For example Grenet (2009) shows that mothers from different professional occupations have very different probabilities of having children across the year. The author solves this problem by controlling for parental characteristics in the IV regression which was not available in early work. We have a large amount of information on students and their parents and therefore include all the *X* controls available in equations (4) and (5) to counter any possible non randomness of timing of birth. Second there has been a lot of discussion on the problems of early grade retention and tracking. The issue is that the age of students at the time they are observed in a class cohort may already reflect the impact of timing of birth on their previous achievement (e.g. repeaters or skippers). Bedard and Dhuey (2006) propose to sort this problem by using 'assigned relative age' of pupils to overcome these methodological issues. This is simply a relative measure of student age measured as $z_i = 12 - m_i$ where m is the month of birth of individual i. We generate results using z_i instead of *QBirth*, in equations (4) and (5) which will give us a monthly estimate of the impact of month of birth on educational attainment, net of grade repetition. Finally we will include school level fixed effects to our models to obtain IV estimates of the causal relationship of schooling on crime which account for all unobserved school specific characteristics. We must still stay conservative in our interpretation of our results because of the potential remaining biases of our instrument.

3- Data and Descriptives

For our analysis we will use two main datasets which were matched to form a uniquely wealthy source of information on young individuals. First we have access to the Secondary Education Student Cohort 1989 (VOCL99) collected by Statistics Netherlands (CBS). This is a large representative longitudinal survey of 17,774 Dutch students from a random sample of 126 schools who were in the first grade of secondary school in 1999. The educational career of these pupils was followed up until 2006 making it possible to know precisely the level of qualification they obtained by age 19. Furthermore, tests of school performance and non-verbal intelligence were administered in the first of their secondary education. A written questionnaire was also given to the parents of the students at the start of the survey with the aim of collecting information about the families and the pupils. This provides us with a very large number of observed characteristics of the youths in out cohort and their parental background which will be used as the *X* controls we described in our modeling strategies. Our second data source comes from register data on every individual arrested by the police on suspicion of committing a crime

between 1999 and 2006 in the Netherlands[1]. This dataset, the Suspects Identification System (HKS) is recorded yearly and was matched to the VOCL99 to both the students in this cohort *and* their parents. We consequently have chronological information on educational career and arrest frequency for our sampled pupils from the start of secondary school aged 12 on average until the end of their teens.

The definition of being 'criminally active' is relatively straight forward and will be 1 if a student appears in the HKS and 0 otherwise[2]. We will however differentiate between being suspected of a crime before the age of legally being able to choose to drop out of school (12 to 16) and later arrests which will be our outcome variable of interest. This enables us to control for the effect of early criminal involvement when estimating the impact of educational attainment on being arrested aged 16 to 19. To measure low educational achievement we use the 1993 Dutch Ministry of Education definition of the minimum level of education a person should acquire to be successful on the labor market: the 'starter's qualification' (*startkwalificatie* in Dutch). Those who leave school without a diploma of upper secondary vocational education (MBO) or upper secondary general education (HAVO or VWO) are considered to be early school-leavers. Although this concept does not amount to a diploma of finishing secondary education per-se, it is a widely used for political and research purposes to distinguish youths with low educational achievement. In our analysis we will therefore aim to measure the impact on criminal activity of individual who obtain a starter's qualification (*Qual = 1*) against their cohort peers who drop out of school before doing so (*Qual = 0*).

Table 1 reports a large number of observable characteristics which we have available from our dataset for the students with and without educational qualification. The last column reports raw differences in these characteristics and if they are statistically significant. We immediately see that these two groups are very heterogeneous and for example those with no qualification are 13 percentage points more likely to be arrested aged 16 to 19. As this could be explained by differences in other individual characteristics which are reported in Table 1, we will want to consider how this probability changes when we include them in our different estimation strategies later. In Figure 1 we report the year and month of birth in our cohort. We can see that the majority of students are born between October 1986 and September 1987 and they represent 'normal' age distribution to enter secondary school. The substantial number of pupils are also

[1] Note that this is a dataset of suspected individuals who are not all convicted of the crime they are arrested for. Still more than 90 percent of individuals are at a later stage found or plead guilty to the crime they were accused of.

[2] One may wonder if arrest is a good indicator of criminal participation. This would be especially problematic if smarter or more educated pupils are less likely to be caught for the crimes they commit. Evidence from self report data in the US (Lochner 2004, Lochner and Moretti 2004) suggests that this is not the case and that education does not affect the probability of arrests.

born before that are likely repeated a grade in primary school and those born after to have skipped a grade or entered education early. Figure 2 shows the age the pupils of the VOCL99 when they were arrested for the first time by the police. The graph clearly shows the age-crime distribution with yearly increases in the number suspected of a crime until age 19. The drop in the final year is only due to there being fewer students who have turned 20 by 2006, the last year we have HKS data for. We therefore will only consider arrests until age 19 to insure that the age effect does not impact on probability of arrest because of this phenomenon.

Since our instrument for educational achievement will be timing of birth, we consider the randomness of its distribution for the individuals in our sample. Figure 3 shows the variation from a normal monthly distribution of birth for the students in the VOCL99. We see that there is relatively little difference in this distribution except perhaps in the final three months of the year. In Figure 4 we check if this is not a 'natural' occurrence due to seasonal variation in the overall Dutch population in birth planning decision. We do see that birth for all boys and girls born in 1986 and 1987 in the Netherlands are also relatively lower in the last three months of the year but not to the magnitude observed in out cohort. There remains the possibility that the observe variation results from the difference in the probability of repeating a grade in primary school driven by timing of births. When we split our sample between repeaters and non repeaters in Figure 5, we do find that there is a large difference from the normal distribution for the former group. This confirms the importance of taking into account observed relative age when we carry out our IV estimation of the causal impact of educational attainment on criminal participation. We now turn to the results from our various estimation strategies we described in the previous section.

4 – Results

We first implement a probit estimation of the link between obtaining a starter's qualification and arrest probability building up the model to include and increasing number of individual observed characteristics as control. Table 2 reports the results from this strategy with the first column reporting β, the raw correlation between *Qual* and *Arrest* which still stands at 13 percentage points. Simply controlling for gender and age in column (3) already reduces this coefficient by almost third. When also including ethnicity, parental crime, and earlier criminal participation in column (6), we find that having some secondary education qualification reduces the probability of arrest by about 7 percent. Early criminal participation appears here as one of the most important predictor of the likelihood of future arrest. In the last column of Table 2 we now

include the full set of individual observable controls we have available[3]. The estimated β is now one percentage point lower with secondary education qualification linked to a 5 percent decrease in the probability of arrest aged 16 to 19. This is much lower than the raw correlation we originally estimated and suggests that schooling has little impact on criminal activity once we control for a large number of individual characteristics. The coefficients on each of these characteristics are reported in the first column of Table 3. The second column shows results once we also include school level fixed effects in the model to control for unobserved educational establishment specific characteristics. Although this changes the value of the coefficients of some of the covariates we include in the model, it does not at all change the estimate on the impact of schooling on probability of arrest. Whether coming out of the same school or at the national level, the impact of obtaining a secondary degree on criminal participation is the same.

In Table 3 we report the estimates from our PSM estimation strategy. Panel B shows the marginal effects of the individual characteristics on the probability of finishing secondary education used to generate propensity scores to match students. Figure 6 illustrates the distribution of these propensity scores for the students who obtain a starter's qualification and those who do not. Still there are a large enough number of individuals with similar scores to insure that the matching will be successful with enough 'common support' from the treated and control groups. The distribution of the later group is clearly skewed to the left showing again showing how different they are in observable characteristics to their more educated peers Two interesting results are that ethnicity and repeating a grade in primary school do not seem to impact on obtaining a starter's qualification in Table 3. However being arrested aged 12 to 15 strongly decreases your chances of getting a secondary education degree emphasising the complexity of the school-crime relationship. The estimated βs in Panel A are slightly larger now but not statistically different than with our previous probit model. We therefore conclude that secondary education completion reduces probability of arrest likelihood by 5 percentage points using the best probability models. Since these methodologies may not account for endogeneity and unobservable characteristic issues of the school enrolment and criminal participation decision relationship, we now turn to our IV estimation strategy

Table 5 reports the first stage, Panel B, and reduced form, Panel A, estimates from the 2SLS models. Columns (1) and (2) use quarter as and instrument and columns (3) and (4) the assigned

[3] Certain questions in the survey were not answered by a non-negligible number of students and parents in the survey. We report in Table A1 of the Appendix the mean distribution of the observed characteristics for the individuals who did not answer these questions. As we observe that these are different to the average distribution of the whole sample, we include dummy variable for each missing observations to capture the effect they may have on our estimates.

relative age. In both cases the first stage estimates of the impact of timing of birth on educational attainment appear very significant with relatively older students less likely to complete secondary school with a valid qualification. When we include school level fixed effects in the last column, we find that students born 12 months earlier than their classmates are roughly an 8 percent lower chance of obtaining a starter's qualification. In the reduced form this is translated into an approximate 34 percent lower probability of arrest resulting from completing a recognised high school degree. Even if we are very conservative with this estimate (because of the remaining identification issues of the instrument mentioned earlier) and consider its lower bound, we still find that a starter's qualification reduces arrest probability by more than 11 percentage points. This suggests that schooling has a much stronger impact on criminal activity, at least twice larger, than models which do not consider the contemporaneous education and offending decision making problematic would suggest. Intuitively one can understand this result as the effect of forcing a 'drop out' to stay on and obtain a starters' qualification on arrest probability. We conclude that this effect is very large when measure accurately and believe it should be considered when designing policies to retain students in school until they reach a minimum level of educational qualification.

5 – Conclusion

The relationship between schooling and crime is a complex one. There is a clear link between educational attainment and arrest probability observed for adults as well as young offenders. Theory suggests that this is driven by the poor present and future wage expectations of individuals with low education qualification. Investigation into the contemporaneous school and crime decision making among youths is complicated by unobservable characteristics which are likely to impact on both choices simultaneously. Moreover being criminally active while at school may reduce your chance of obtaining qualification which may in turn increase your probability of crime participation. The school-crime relationship is plagued by these endogeneity issues which make it difficult to identify and measure the causal relationship. In this paper we suggest an identification strategy which instruments educational attainment by timing of birth of pupils. We argue that this should have little impact on crime except through its effect on differences in school qualification. This IV approach should enable us to accurately measure the direction and size of the impact of educational attainment on crime.

Using a large survey of Dutch youths entering secondary school matched to their criminal history, we obtain a number of estimates of the link between school performance and arrest probability. Our probit estimates show that the link between obtaining a qualification on crime participation greatly reduced once individual characteristics are included in the models. Obtaining a starter's qualification is linked to a 5 percentage point decrease in arrest probability and this result is robust to including school level fixed effects. Propensity score matching also yields similar estimates suggesting that school has at first an apparently small impact on criminal participation. When instrumenting educational attainment by observed relative age we now find that it reduces the chances of arrest by more than 30 percent. Even the most conservative interpretation of these results leads us to conclude that obtaining a secondary education qualification is twice more effective in reducing crime participation then when estimated with models who do not deal with the endogeneity issues it raises.

We believe this to be a more realistic measure of the school crime relationship which should be considered as an important factor to promote policies that incentivise students to stay on at school. This is also because youth will tend to make an early choice between little education and a life of street crime or a good education and a largely crime-free life. Forcing some young individuals to obtain a minimum level of qualification could therefore have long term effect by decreasing the probability of a criminal career which becomes much more difficult to interrupt later in life.

References

Angrist, J. and A. Krueger (1991). Does Compulsory Schooling Attendance Affect Schooling and Earnings?, *Quarterly Journal of Economics*, 106, 979-1014

Angrist, J. and A. Krueger (1992).The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples., *Journal of the American Statistical Association*, 87, 328-336

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

Crawford, C., Dearden L., and C. Meghir (2007). When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England., *Institute of Fiscal Studies External Report*, IFS, London.

Freeman, R. (1996). Why Do So Many Young American Men Commit Crime and What Can We Do About It?, *Journal of Economic Perspectives*, 10, 25-42.

Freeman, R. (1999) The Economics of Crime., in… (Eds), *Handbook of Labor Economics*,

Gould, E., Mustard, D., and B. Weinberg (2002). Crime Rates and Local Labor Market Opportunities in the United States: 1977-1997., *Review of Economics and Statistics*, 84, 45-61.

Grenet, J. (2009) Academic Performance, Educational Trajectories and the Persistence of Date of Birth Effects. Evidence from France. *Mimeo*, CEP, London School of Economics

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

Hjalmarsson, R. (2008) Criminal Justice Involvement and High School Completion., *Journal of Urban Economics*, 63, 613-630

Travis, H. and M. Gottfredson (1983). Age and the Explanation of Crime., *American Journal of Sociology*, 89, 552-584.

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

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

Lochner, L. (2009). Education and Crime., in McGaw B., Baker E., and P.P. Peterson (Eds), *International Encyclopaedia of Education, 3rd Edition*, Oxford: Elsevier, in press.

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

Machin, S., and Meghir, C. (2004). Crime and Economic Incentives., *The Journal of Human Resources*, 39, 958-979.

Ministerie van OC&W (1993) *Een goed voorbereide start*, Den Haag.

Rosenbaum, P. and D. Rubin (1983) The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika*, 70, 41-55.

Rosenbaum, P. and D. Rubin (1984) Reducing Bias in Observational Studies Using Subclassification on the Propensity Score, *Journal of the American Statistical Association*, 79, 516-524.

Figure 1: Month and Year of Birth of VOCL99 Cohort

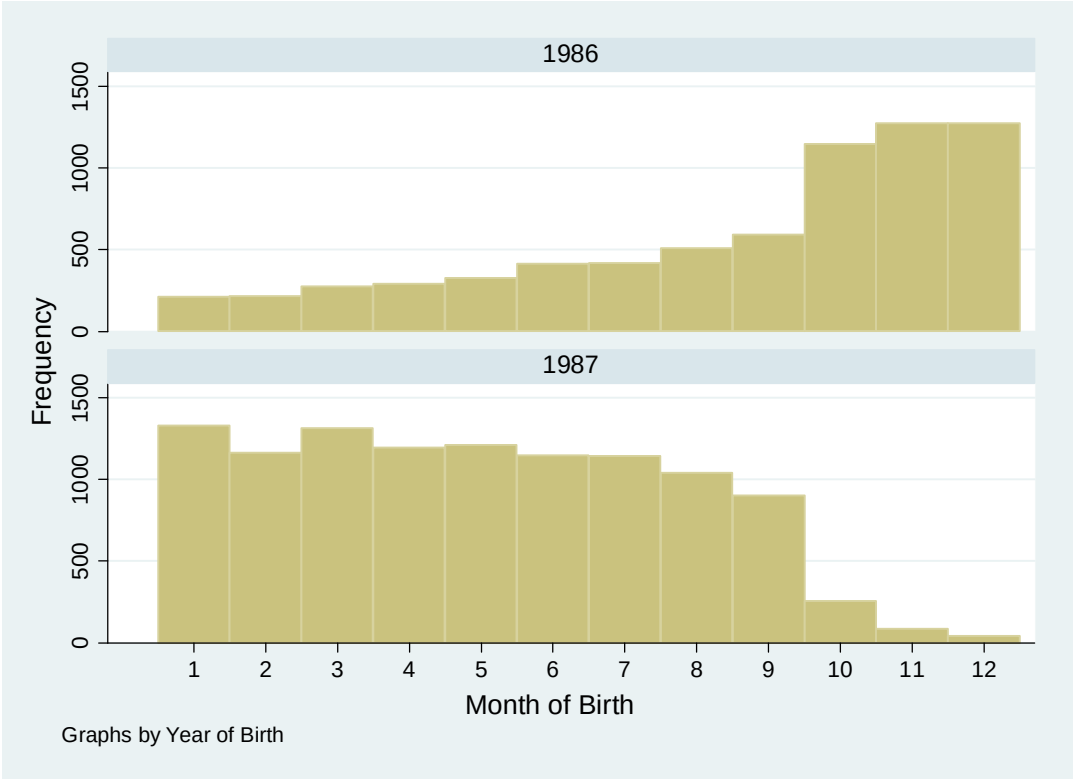

Figure 2: Age Distribution when Suspected of First Crime

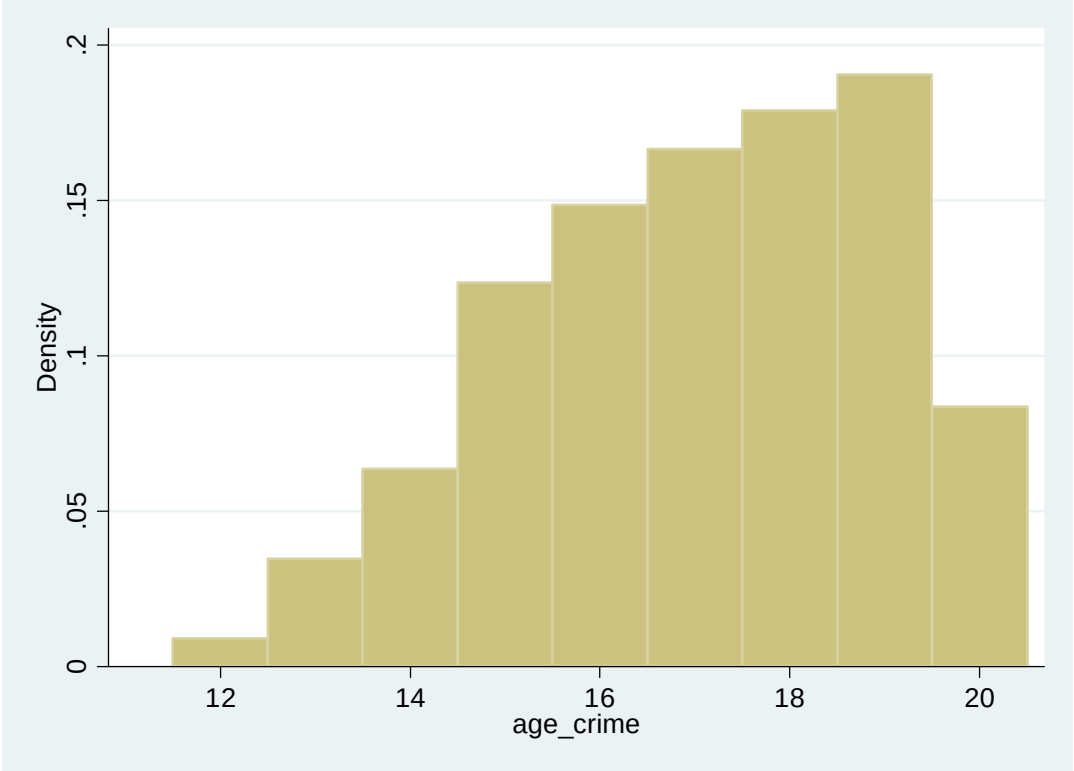

Figure 3: Difference in Month of Birth Distribution in VOCL Cohort

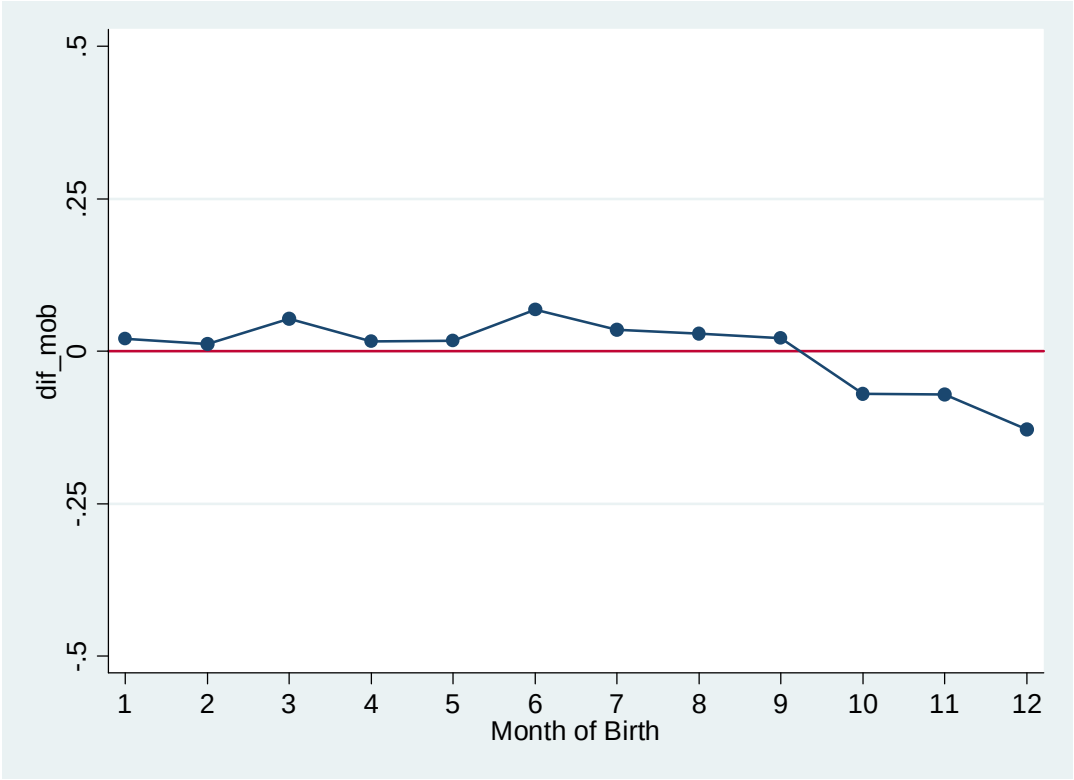

Figure 4: Difference in Month of Birth Distribution – Dutch Births 1974-1999

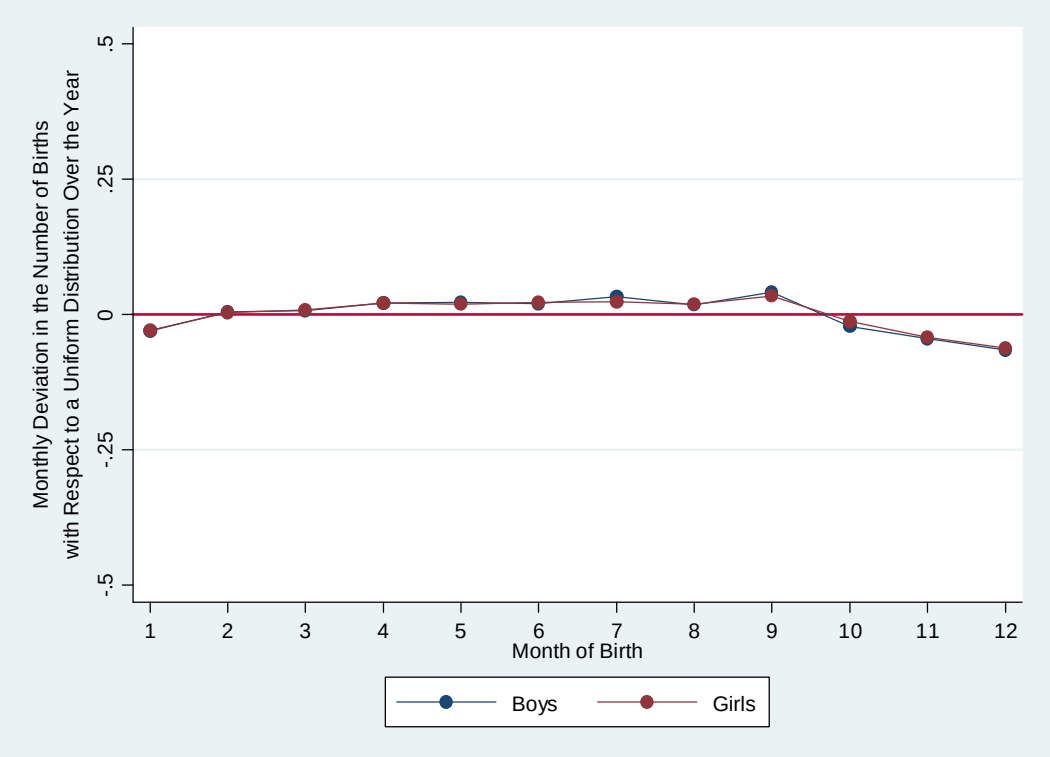

Figure 5: Difference in Month of Birth Distribution – Repeaters

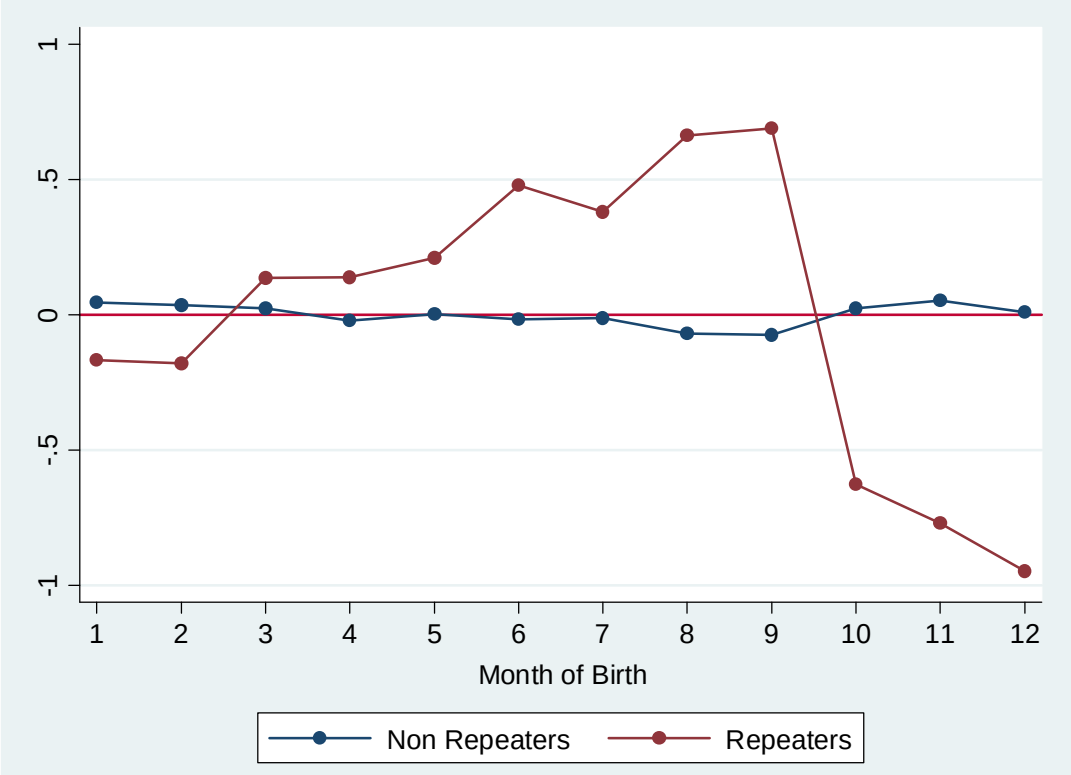

Figure 6: Propensity Score Distribution – No Vs Some Education

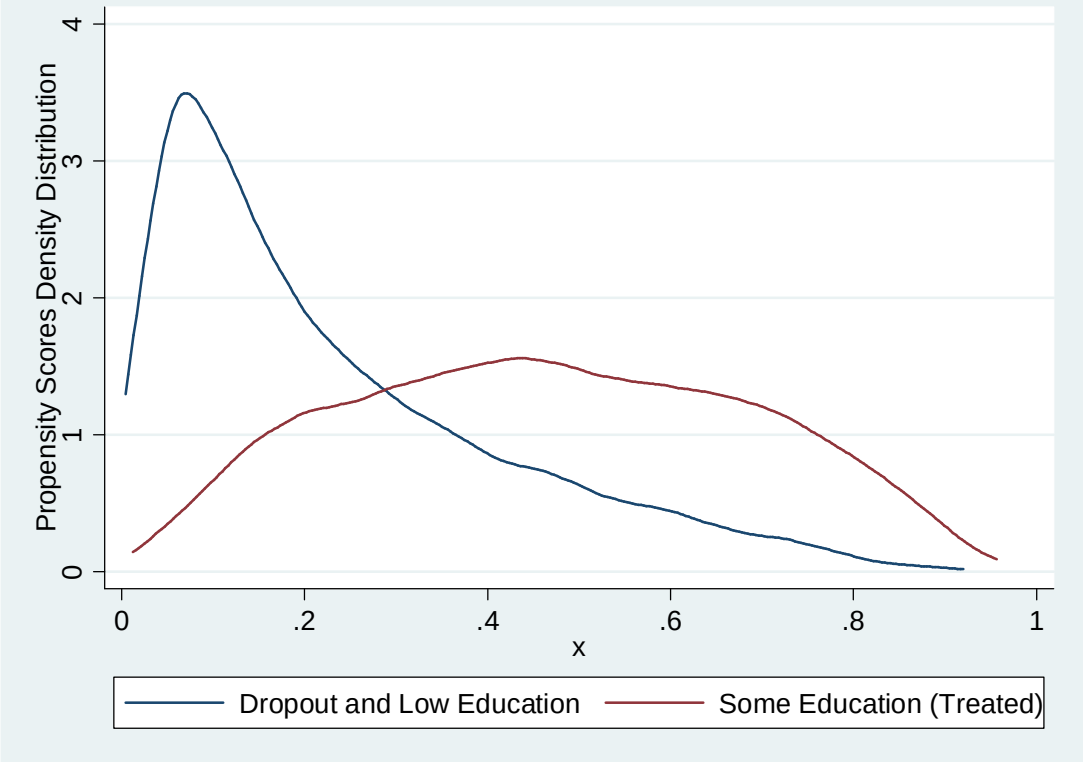

Table 1: Descriptive Statistics by Education Attainment Levels of Individuals

	No Vs Some Education Qualification		
	No	**Some**	**Difference**
Crime Age 16-19	.172	.043	-.129*** (.004)
Gender	.558	.469	-.089*** (.008)
Age	12.65	12.48	-.173*** (.007)
Ethnicity	.242	.154	-.088*** (.006)
Parental Crime	.043	.016	-.027*** (.002)
Crime 12 to 15	.087	.018	-.068*** (.003)
Repeated Primary	.394	.182	-.212*** (.007)
Mother Age	40.76	41.68	.920*** (.072)
Parental Income	10.02	10.25	.236*** (.016)
Religious Parents	.678	.684	.006 (.007)
Married Parents	.746	.853	.107*** (.006)
Parental Education	12.36	14.20	1.84*** (.052)
Ability Test at Age 12	.339	.583	.243*** (.004)
Pupil School Motivation	.481	.508	.027*** (.005)
Pupil School Perception	.486	.509	.022*** (.005)
Parental Schoolwork	.556	.477	-.079*** (.004)
Parental Interest in School	.474	.515	.041*** (.004)
Parental Cultural Capital	.494	.500	.007 (.004)
Parental Authority	.272	.292	.019*** (.003)
Sample Size	5,408	12,366	-

Table 2: Probit Results – Building Up Model

	Dependent Variables = Suspected of a Crime Aged 16-19						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Some Qualification	-.129*** (.005)	-.109*** (.005)	-.098*** (.005)	-.092*** (.005)	-.090*** (.005)	-.073*** (.005)	-.049*** (.004)
Gender	-	.096*** (.004)	.092*** (.004)	.093*** (.004)	.093*** (.004)	.084*** (.004)	.082*** (.004)
Age	-	-	.032*** (.004)	.028*** (.004)	.028*** (.004)	.032*** (.004)	.026*** (.004)
Ethnicity	-	-	-	.042*** (.005)	.041*** (.005)	.030*** (.005)	.018*** (.004)
Parental Crime	-	-	-	-	.071*** (.016)	.053*** (.014)	.035*** (.012)
Crime Aged 12-15	-	-	-	-	-	.218*** (.017)	.183*** (.016)
All Other Controls	No	No	No	No	No	No	Yes
Missing Dummies	No	No	No	No	No	No	Yes
Observations	17,774	17,774	17,774	17,774	17,774	17,774	17,774

Table 3: Probit Results – Without and With School Fixed Effects

Dependent Variables = Suspected of a Crime Aged 16 to 19		
	(1)	**(2)**
Some Qualification	-.049*** (.004)	-.049*** (.004)
Gender	.082*** (.004)	.079*** (.004)
Age	.026*** (.004)	.025*** (.004)
Ethnicity	.018*** (.004)	.015*** (.004)
Parental Crime	.035*** (.012)	034*** (.014)
Crime 12 to 15	.183*** (.016)	175*** (.016)
Repeated Primary	-.008 (.005)	-.008* (.005)
Mother Age	-.001* (.000)	-.001** (.000)
Parental Income	-.000 (.002)	-.001 (.001)
Religious Parents	-.014*** (.003)	-.010*** (.004)
Married Parents	-.013** (.004)	-.010** (.004)
Parental Education	-.002*** (.000)	-.002*** (.000)
Ability Test at 12	-.039*** (.006)	-.044*** (.007)
School Motivation	-.006 (.006)	-.006 (.005)
School Perception	-.001 (.005)	.000 (.006)
Parental Schoolwork	.014** (.006)	.013** (.006)
Parental Interest	-.015** (.006)	-.015*** (.006)
Cultural Capital	-.003 (.006)	-.001 (.006)
Parental Authority	.004 (.007)	-.007 (.007)
School Fixed Effects	No	Yes
Missing Dummies	Yes	Yes
Observations	17,774	17,727

Table 4: PSM Results – With and Without Fixed Effects

Panel A: Dependent Variables = Suspected of a Crime Aged 16 to 19		
Education	-.045*** (.012)	-.061*** (.012)
Panel B: Marginal Effects from Probit Regression: Pr(Education Level = 1)		
Gender	-.056*** (.007)	-.059*** (.007)
Age	-.083*** (.010)	-.079*** (.010)
Ethnicity	.001 (.010)	-.011 (.010)
Parental Crime	-.084*** (.025)	-.085*** (.025)
Crime 12-15	-.222*** (.022)	-.229*** (.023)
Repeated Primary	-.013 (.013)	-.000 (.013)
Mother Age	.004*** (.001)	.003*** (.001)
Parental Income	.018*** (.003)	.017*** (.003)
Parents Religious	.018** (.008)	.035*** (.009)
Parents Married	.070*** (.011)	.074*** (.012)
Parental Education	.014*** (.001)	.012*** (.001)
Ability Test at 12	.543*** (.014)	.462*** (.017)
School Motivation	.055*** (.013)	.038*** (.014)
School Perception	-.006 (.013)	-.008 (.013)
Parental Schoolwork	-.188*** (.015)	-.200*** (.015)
Parental Interest	.126*** (.015)	.117*** (.015)
Cultural Capital	.028* (.014)	.020 (.015)
Parental Authority	-.005 (.017)	-.001 (.017)
School Fixed Effects	No	Yes
Missing Dummies	Yes	Yes
Observations	17,774	17,774

Table 5: IV Results – Education Attainment Instrumented by Quarter of Birth or Relative Age. With and Without School Level Fixed Effects

	Dependent Variables = Suspected of a Crime Aged 12 to 19			
	Dropout and Low Vs Some Qualification			
	(1)	**(2)**	**(3)**	**(4)**
Reduced Form				
Some Qualification	-.365*** (.115)	-.397*** (.131)	-.300*** (.093)	-.341*** (.115)
IV First Stages: Education =Quarter of Birth *or* Education = Relative Age in Months				
Relative Age	-	-	-.008*** (.001)	-.006*** (.001)
January-March	-.070*** (.012)	-.064*** (.012)	-	-
April-June	-.081*** (.012)	-.045*** (.012)	-	-
July-September	-.034*** (.012)	-.032*** (.012)	-	-
Controls	Yes	Yes	Yes	Yes
School Fixed Effects	No	Yes	No	Yes
Observations	17,774	17,774	17,774	17,774

Note: Robust standard errors in parenthesis.

Appendix

Table A1: Sample Selection of Missing Information

	Means for the Following Individual Characteristics						
Missing Information	**Male**	**Age**	**Ethnic**	**No Qual**	**Crime**	**Parent Crime**	**N**
0. None	49.61	12.53	18.07	30.42	10.34	2.40	**17,774**
1. Ability Test	51.10	12.60	24.50	41.91	14.83	3.64	**1,045**
2. School Perception	52.75	12.59	24.53	41.51	16.18	3.77	**1,378**
3. Repeated Primary	54.81	12.62	36.83	46.10	18.94	4.71	**1,974**
4. Parental Religion/Marriage	55.00	12.61	37.29	46.44	19.20	4.78	**1,880**
5. Parental Education	54.19	12.61	35.86	45.57	18.89	4.45	**2,181**
6. Parenting Style	53.34	12.60	30.71	42.64	16.80	4.19	**2,934**
7. Cultural Capital	54.76	12.61	39.00	46.43	19.92	5.11	**1,682**
8. Social Capital	54.14	12.61	34.66	45.37	18.11	4.43	**2,259**
9. Siblings and Rank	53.92	12.61	35.80	45.15	18.43	4.45	**2,268**
10. Peers	55.49	12.56	25.50	38.13	15.99	3.15	**2,701**
11. Social Big Five	52.34	12.57	20.30	36.10	13.10	2.80	**5,395**