

The effect of childhood conduct disorder on human capital

Dinand Webbink*, Sunčica Vujić**, Pierre Koning*** and Nicholas G. Martin****

Abstract

This paper estimates the longer-term effects of childhood conduct disorder on human capital accumulation and violent and criminal behaviour later in life using data of Australian twins. We measure conduct disorder with a rich set of indicators based on diagnostic criteria from psychiatry. Using ordinary least squares (OLS) and twin fixed effects (FE) estimation approaches, we find that early (pre-18) conduct disorder problems significantly affect both human capital accumulation and violent and criminal behaviour over the life course. In addition, we find that conduct disorder is more deleterious if these behaviours occur earlier in life.

Key words: conduct disorder, human capital, twins

JEL code: I1, I2, K42

* Erasmus School of Economics Rotterdam, Tinbergen Institute, IZA

** London School of Economics and Political Science

*** CPB Netherlands Bureau for Economic Policy Analysis, IZA

**** Queensland Institute of Medical Research, Brisbane

Corresponding author: webbink@ese.eur.nl

1 Introduction

Many children have mental health problems which hinder their normal development and functioning. Anxiety, conduct, attention, and depressive disorders are the most common. For example, up to 500,000 (14%) of Australian children between the ages of 4 and 18 years have significant mental health problems (Australian Centre for Community Child Health, 2006); up to 50% of 'hard to manage' pre-school children are likely to have continuing hardships requiring professional help; approximately one in five children and adolescents in the U.S. may have a mental health disorder (Currie & Stabile, 2006). Despite these large numbers, little is known on the longer-term effects of these mental health problems of children. Currie & Stabile (2009) note that most studies 'assume that early mental health problem will have negative effects and focus on the efficacy of specific interventions'.

This paper analyses the longer-term effects of childhood conduct disorder on human capital accumulation and violent and criminal behaviour later in life using data of Australian twins. If someone shows "a repetitive and persistent pattern of behaviour in which the basic rights of others or major age-appropriate societal norms or rules are violated," then he/she falls into a category of individuals with a conduct disorder problem. Conduct disorder is known as a 'disruptive behaviour disorder' because of its impact on children and their families, neighbours, and schools, and is largely associated with delinquent or criminal activity. We measure conduct disorder using diagnostic criteria from the American Psychiatric Association (APA, 1994). In line with Currie & Stabile (2007) we estimate effects on positive human capital, including measures such as grade repetition, marks in high school and high school graduation, and on 'negative' human capital, including measures such as being arrested, spent time in jail and physically attacking others.

The two main challenges in estimating the longer-term effects of childhood conduct disorder on human capital are the measurement of conduct disorder and omitted variable bias. This paper contributes to the economic literature by addressing these two main challenges. First, it is easy to know when a child has a fever but a child's mental health problem is harder to identify. Typically, mental health problems in children are diagnosed by asking a child's parents and teachers a series of questions about their behaviours. For instance, Currie and Stabile (2009) use 6 questions to form a conduct disorder scale. In this study, we can employ a much richer set of conduct disorder indicators. We use self-reports of adult twins on 21 statements that follow the definition of conduct disorder according to the APA criteria. In addition, the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) emphasizes that there should be at least three specific conduct disorder behaviours present within the same 12 month period to make the diagnosis of conduct disorder (APA, 1994). Our data include self-reports on a question that exactly matches this definition. Moreover, we have information whether the twin saw a doctor, psychologist or other professional for these behaviours. We used this information for constructing three measures of conduct disorder. Another contribution of this study is that we

also have information on the age at which the conduct disorder behaviours occurred. We use this information to investigate the effects of the timing of the problem behaviours on human capital.

Second, estimates of the effect of conduct disorder on human capital might be confounded by unobserved differences between children and their families. For example, in Australia the incidence of mental health problems is even higher in disadvantaged children, such as Aboriginal children (24%), children residing in 'out of home care' (55-60%) and children with a disability, who are up to four times more likely to have mental health problems than children without a disability (Australian Centre for Community Child Health, 2006). Parental substance abuse, marital conflict, psychiatric illness and child abuse and neglect have been identified as risk factors for conduct disorder (Searight, et al., 2001). Previous studies on the effects of childhood mental health problems mainly used cross sectional and within-sibling estimation (Currie and Stabile, 2006, 2007; Fletcher and Wolfe, 2008).

In this paper we estimate within-family models using data of (identical) twins. We are aware of only one previous study (Le et al. 2005) that also uses data on twins. The advantage of using twins instead of siblings is that the family circumstances for twins will typically be more similar than with siblings. More importantly, identical twins are genetically identical, whereas siblings on average only share half of their genetic endowments. Therefore, using data on twins, in particular those that are genetically identical, may reduce the bias caused by heterogeneity within families. In addition, we can further reduce this bias by controlling for differences in birth weight within pairs of identical twins. Recent research has shown that birth weight is an important predictor of later outcomes in life (Black et al., 2007).

In the economic literature, several recent studies investigate the longer-term effects of mental health problems on human capital. Most papers find that mental disorders, and especially ADHD, have large negative effects on human capital accumulation. Currie and Stabile (2006) examine the relationship between hyperactivity (ADHD) symptoms and short-term human capital outcomes (test scores, grade repetition, special education, and delinquency). Using OLS and within sibling estimation they find large negative effects on test scores and schooling attainment. Fletcher and Wolfe (2008) estimate the effect of ADHD on human capital. Their OLS results imply that children with ADHD face longer-term educational disadvantages. Once family fixed effects are controlled for, standard errors become larger, thus rendering insignificant estimation results. However, they find evidence for spillover effects within families. They also investigated the effect of ADHD on self reported crime (Fletcher and Wolfe, 2007). Controlling for sibling fixed effects, they find that both inattentive and hyperactive symptoms during childhood increase the likelihood of criminal behaviour. Currie and Stabile (2009) address the effects of depression and conduct disorder (antisocial behaviour/aggression). For the US, they find that conduct disorder has negative effects on various human capital outcomes, while for Canada they find only negative effects on the probability that 16-19 year old youths are in school (fixed effect estimates for the other

outcomes suggest negative effects but are statistically insignificant). Le et al. (2005) investigated the effect of childhood conduct disorder on early school leaving and labour market outcomes, such as employment and earnings, in Australia. Their findings suggest that individuals who experienced conduct disorder problems are more likely to leave school early, have poorer employment prospects and lower earnings. However, the within-twin estimates of the effect of conduct disorder on early school leaving conducted on the sample of identical twins are statistically not significant. They therefore conclude that genetic factors are responsible for the positive link between conduct disorder and early school leaving.

In this paper we use the same data as Le et al. (2005) and also investigate the impact of conduct disorder on early school leaving (high school graduation). However, we address a broader range of outcomes. In addition to high school graduation, which can be seen as an indicator of the quantity of human capital, we also investigate the effects on grade repetition and marks in high school. These variables are related to other dimensions of human capital accumulation such as the time needed to attain a certain education level and the quality of the performance in high school. Moreover, we also focus on a set of indicators of negative human capital, such as aggressive and criminal behaviour. This might be important because conduct disorder is largely associated with delinquent or criminal activity. Finally, by using three different measures, instead of one measure, of conduct disorder (see next section) we might improve the measurement of conduct disorder. This seems important as mental disorders typically are difficult to observe.

We find deleterious effects of conduct disorder on positive and especially on negative human capital, even within pairs of identical twins. For instance, within pairs of identical twins we find that conduct disorder reduces the probability of being arrested by 7 to 16 percent points. Another important finding is that the effect of conduct disorder on human capital is more deleterious if these behaviours occur earlier in life. A wide range of robustness checks suggest that our main estimates may be lower bounds of the true effect of conduct disorder on human capital. We conclude that childhood mental health problems have high human and financial costs for families and society at large.

The remainder of this paper is organised as follows. The next section describes the data. Section 3 explains the methodology. Section 4 shows the main estimation results. Robustness checks are presented in Section 5. Section 6 concludes and discusses some policy implications.

2 Data

In this study, we analyze data from the so-called younger cohort of twins of the Australian Twin Register (ATR) which were gathered in two surveys, in 1989-1990 and in 1996-2000. The surveys gathered information on the respondent's family background (parents, siblings, marital status, and children), socioeconomic status (education, employment status, and income), health behaviour (body size, smoking and drinking habits), personality, feelings and attitudes. More details about the collection of the data can be found in Le et al. (2005) and in appendix A.1 in this paper.

Conduct disorder

The measurement of conduct disorder is crucial for our analysis. The American Psychiatric Association (APA) has defined a set of criteria for the diagnosis of conduct disorder such as: often initiated physical fights; has deliberately destroyed others' property; has broken into someone else's house, building, or a car; has often been truant from school, etc. (for a full definition see Table A.1 in the Appendix). Our data contains self-reported information (when twins were on average 30 years old) on 21 statements which reflect these behavioural problems before the age of 18.

In the empirical analysis, we will use three measures of conduct disorder:

1. *Conduct disorder score* sums up occurrences of 21 behavioural statements above;
2. *APA Definition* is based on the question 'Did you do at least 3 of these things within the same 12-month period?';
 - a. We use answer to the question 'How old were you the first time you did at least 3 of these things within the same 12-month period?' to investigate the effect of the timing of conduct disorder.
3. *Professional help* is based on the question 'Did you ever see a doctor, psychologist, social worker or any other professional like that because of these behaviours?'

The psychiatric handbook (DSM-IV) emphasizes that at least three conduct disorder behaviors should occur within the same 12-month period to make the diagnosis of conduct disorder. Since 'APA Definition' perfectly matches the definition of conduct disorder according to the American Psychiatric Association, we will use it as our main measure in the analysis. For a more detailed description on how these measures of conduct disorder are constructed, see Appendix A.2.

Positive and negative human capital

In line with Currie and Stabile (2007), we distinguish between positive and negative human capital. Positive human capital is the type of human capital that is generally accumulated through schooling. Negative human capital is related with criminal, antisocial and violent activities. We use three measures of positive human capital: grade repetition, marks in high school and high school graduation. Marks in high school are measured with a three-point scale (below average, average, better than average). High school graduation is a dummy variable for completing at least 11.5 years of education. We also use three measures of 'negative' human capital: 'physically attacking others'; 'being arrested since you were 18'; 'spent time in jail'. All three negative human capital variables are dichotomous variables. Unfortunately, due to the routing of the questionnaire twins with a conduct disorder score of zero did not answer several questions on human capital. As this may bias the estimates downward we checked the sensitivity of the results by imputing mean values on these outcomes for twins with no childhood conduct disorder (see section 6). In the analysis we use as covariates: mother's and father's education, age, age squared, gender and birth weight.

The main variables in our analysis (conduct disorder, positive and negative human capital) are all based on self-reports. The reliability of these self-report data is an important issue. In this paper we use three different measures of conduct disorder and one of our measures is based on 21 statements. We expect that the use of such a rich set of indicators improves the reliability of the data. Self-reports on conduct disorder have been shown to have acceptable reliability (Slutsky et al. (1997)). In criminology, a large literature shows that self-report data have consistently acceptable reliability and validity. Many studies find high correlations of self-report data with other criterion related measures of criminal frequency and arrest histories (Farrington, 1973; Hardt & Hardt, 1977; Horney & Marshall, 1992; Huizinga & Elliott, 1986; Maddux & Desmond, 1975; Mieczkowski, 1990; Weiss, 1998). Thornberry and Krohn (2000) conclude that "self-reported measures of delinquency are as reliable as, if not more reliable than, most social science measures".

Descriptive statistics

Table 2.1 reports means of all measures for pairs of twins with non-missing information on conduct disorder symptoms. Column 1 (3) shows means for complete pairs of all (identical) twins. Column 2 (4) shows the number of twins with a within-family difference in the variable in question for the sample of all (identical) twins. The top panel shows that the average conduct disorder score is nearly 2. Currie and Stabile (2007) report averages of nearly 5 for the US and 1.5 for Canada using a scale from 0 to 16. Hence, our Australian sample seems more comparable to the Canadian sample than to the US sample. The second row of table 3.1 shows that 13 percent of our sample would be diagnosed as having conduct disorder according to the APA definition, for men and women this is respectively 20 and 8 %. This is roughly in line with

Searight et al. (2001) who report that approximately 6 to 16 percent of boys and 2 to 9 percent of girls meet the diagnostic criteria for conduct disorder. The fraction of our sample that saw professional help is much smaller, approximately 3 percent.

Table 2.1 Means table for sample of children with all conduct disorder measures non-missing

	All twin pairs		Identical twins	
	Mean	Twin differences	Mean	Twin differences
Conduct disorder measures				
Conduct disorder score (1-21)	1.880	3596	1.674	1394
APA definition	0.131	956	0.112	330
Professional help	0.027	264	0.024	92
Positive human capital				
Grade repetition	0.157	792	0.167	250
Marks high school (1-3)	2.3	2186	2.3	748
High school graduation	0.792	1102	0.790	197
Negative human capital				
3+ Attacking others	0.192	658	0.191	252
Arrested since 18	0.074	290	0.069	102
Jail	0.019	70	0.018	22
Covariates				
Education mother	10.3		10.3	
Education father	10.5		10.5	
Age in 1996	29.9		29.9	
Gender (male=1)	0.442		0.409	
Birth weight (grams) N	2540		2430	
N	5322		2250	

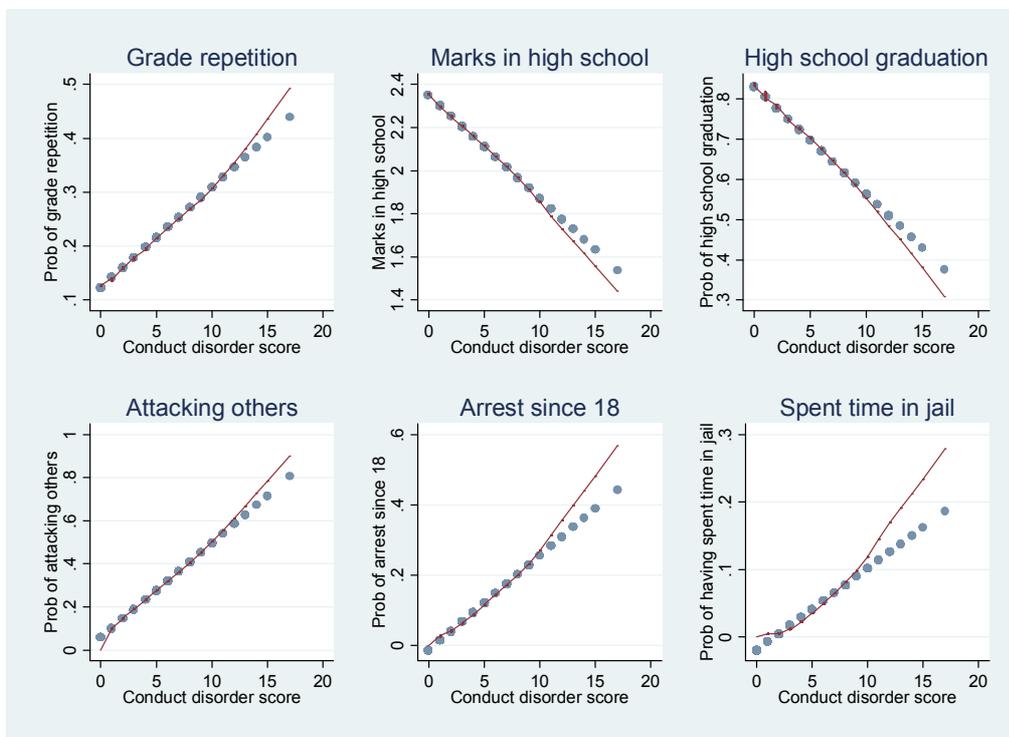
Note: The column 'Twin differences' shows the number of individuals with a within-twin pair difference on a specific variable.

The number of twins with a within-family difference on one of the three measures of conduct disorder is much larger in the sample of all twins than in the sample of identical twins (column 2 and 4). This illustrates that genetic factors, that are exactly the same within pairs of identical twins, are important for the development of conduct disorder. The average (standard deviation) of the within twin differences for the conduct disorder score is 1.9 (2.5) for the sample of all twins and 1.7 (2.3) for the sample of identical twins. The intra class correlation for the conduct disorder score is 0.63 for identical twins and 0.30 for fraternal twins. The smaller amount of variation in the sample of identical twins makes the estimates more vulnerable for measurement error (Grilliches, 1979) and may limit the opportunities for finding effects on human capital. It should also be noted that the number of pairs used in the estimation may be smaller due to missing values for human capital outcomes.

Nearly 39 percent of all twins reported negatively on all 21 statements of conduct disorder and nearly 73 percent have a conduct disorder score of less than three¹. This distribution is remarkably similar to the distribution of conduct disorder (antisocial/aggression) for Canada reported in Currie and Stabile (2007). As expected, the conduct disorder score for twins that reported ‘yes’ on the APA definition or those who sought professional help are much higher, on average more than 5 (4) points higher for the second (third) definition. Behaviours most frequently reported are: wagged school, stay out late, shop lifted, misbehaved at school, stole from home and family.

A first exploration of the relationship between conduct disorder and human capital is shown in figure 2.1. The figure shows the association between the conduct disorder score and our measures of positive and negative human capital.²

Figure 2.1 The associations between conduct disorder scores and human capital



The association between conduct disorder and human capital seems remarkably consistent for all measures of human capital. An increase of conduct disorder is associated with a decrease in positive human capital and increase in negative human capital. For instance, higher levels of conduct disorder are associated with lower marks in high school and a lower probability of graduating from high school. In addition, higher levels of conduct disorder are associated with

¹ These descriptives can be obtained from the authors on request.

² The associations are constructed by locally weighted least squares (LOWESS).

higher probabilities of attacking others, being arrested since the age of 18 or having been incarcerated. In line with previous studies for the US and Canada (Currie and Stabile, 2006, 2007) we find that the association between conduct disorder and human capital seems quite linear. Even low scores of conduct disorder are associated with lower human capital accumulation.

3 Methodology

We follow the same approach as in previous papers on the effects of childhood mental health problems on human capital. First, we estimate OLS models of the relationship between conduct disorder and positive and negative human capital outcomes. As these estimates might be biased by unobserved factors we proceed with estimating within twin fixed effect models:

$$HC_{ij} = \alpha + \beta CD_{ij} + \gamma X_{ij} + f_j + \varepsilon_{ij} \quad (3.1)$$

where HC_{ij} is a measure of positive/negative human capital; CD_{ij} is one of our three measures of conduct disorder; X_{ij} is a vector of covariates, consisting of age, gender, birth weight, and mothers' and fathers education; f_j captures unobserved family effects common to all twins within the same family, and ε_{ij} is a random error term. Index i corresponds to the twin number; index j corresponds to the twin family. In this model the family fixed effect is removed by differencing within pairs of twins. We also use OLS and within-twin models for estimating the effect of the age at which conduct disorder behaviours occur on human capital accumulation. Our data contains information on the age of onset of the behaviours according to the APA definition. We constructed three indicator dummies by slicing up the age-of-onset range into three parts with approximately the same proportion of individuals: 3-12, 13-14 and 15-18. Hence, twins coded as zero on the APA definition have a value of zero on these three dummies. Twins with age of onset of 17 have a value of one on the dummy for the age range 15-18. For investigating the effect of age of onset of conduct disorder we estimated the following fixed effect model:

$$HC_{ij} = \alpha + \beta CD_{ij}^{\leq 12} + \delta CD_{ij}^{13-14} + \lambda CD_{ij}^{15-18} + \gamma X_{ij} + f_j + \varepsilon_{ij} \quad (3.2)$$

It should be noted that the identification of the parameters in this equation mostly relies on within twin pair comparisons of twins with conduct disorder and twins that do not report conduct disorder. As such, the estimates capture both effects due to conduct disorder and the timing of conduct disorder. Controlling for conduct disorder would make it possible to identify both effects. However, in our approach this means that we need pairs of twins who both report positive on conduct disorder but who differ in the timing of conduct disorder. Unfortunately, this is too demanding for the data at hand. Additional to equation (4.2) we also estimated a specification that included years of conduct disorder ($18 - \text{age onset}$) and years of conduct disorder squared.

There are several concerns with estimates based on within-family models. First, the estimates might be biased by within-family heterogeneity. Most previous papers on mental health

problems of children estimate within-family models using data of siblings. However, siblings may differ in genetic endowments. In addition, the socioeconomic conditions facing siblings and the parental inputs received by siblings may differ if family circumstances change over time. As we use data on (identical) twins it seems less likely that our estimates will be biased by within family heterogeneity. The family circumstances for twins will probably be more equal than for siblings, who differ in age. More importantly, identical twins share exactly the same genes whereas siblings on average only share half of their genetic endowments. By estimating fixed effect models for separate samples of identical twins, we control for all differences in genetic endowments. In addition, we test the robustness of the estimates by excluding pairs of twins with very large differences in educational attainment. These large education differences might indicate that these twins are quantitatively different from the rest of the sample and introduce heterogeneity which will confound the effects we are looking for. A potential source of within-family heterogeneity are coexisting conditions such as mood disorders (depression), anxiety, ADHD or substance use disorders. For instance, Currie and Stabile (2007) found that conduct disorder is co-morbid with depression and ADHD. These coexisting conditions might induce omitted variable bias. We address this issue by including controls for a large set of self-reported physical and mental health conditions before the age of 14. These conditions were measured in the first survey. In addition, we control for self-reported depression measured in the second survey. Unfortunately, our data do not include measures of ADHD.

A second concern in within-family models is measurement error. The within-family estimator exacerbates measurement error, which is likely to bias the estimates towards zero (Grilliches, 1979). Unfortunately, our data do not provide a second independent measure of conduct disorder that can be used as an instrumental variable and might solve the problem of measurement error (Ashenfelter, et al., 1994). However, by using three measures of conduct disorder we aim to mitigate this problem. In addition, we investigated recall bias by regressing the APA definition of conduct disorder on age (or age and age squared) controlling for gender, education parents and birth weight. However, all estimates of the age coefficients are statistically insignificant. We also found no evidence for recall bias with respect to the age of onset of conduct disorder.

4 Main estimation results

We examine the effect of conduct disorder problems on different human capital outcomes using three measures of conduct disorder. Table 5.1 reports the estimated effects of conduct disorder on six types of human capital using linear regression (probability) models. The first three columns show the effects of conduct disorder on positive human capital, the last three columns show the estimated effects on negative human capital. The top panel shows the results using the conduct disorder score, the middle panel shows the results for the APA definition of conduct disorder and the bottom panel shows the results using the professional help definition. In each panel we first report OLS estimates, next we report fixed effect estimates for the sample of all twins and finally we report fixed effect estimates for the sample of identical twins.³ All regressions include birth weight, age, age squared, gender, and parents' education as controls. Some of these variables drop out from the fixed effect specifications. Each cell shows the results of a separate estimation.

The estimates in table 4.1 suggest that conduct disorder has a negative effect on human capital accumulation and a positive effect on violent and criminal behaviour. The largest effects are found when using the OLS estimation. As expected, the size of the estimates reduces when twin fixed effects are taken into account. For the full sample of twins all estimates, except for one, remain statistically significant. The standard errors increase when we restrict the sample to identical twins only. However, even for the sample of identical twins, where the variation in conduct disorder is much smaller than in the full sample of twins (see table 3.1), nearly all estimates of conduct disorder on violent and criminal behaviour are statistically significant.

Most remarkable and compelling are the effects of conduct disorder on negative human capital. The estimates of the effects on the probability of being arrested since the age of 18 and on the probability of spending time in jail seem quite large and robust. In addition, the estimates of the effects of conduct disorder on the probability of physically attacking others are very large for the first two measures of conduct disorder.

The estimates of the effects on positive human capital also indicate detrimental effects of conduct disorder. We find negative and significant estimates of conduct disorder on high school graduation in the OLS- and fixed effect regressions using the full sample of twins. However, the effects for the first two definitions of conduct disorder become statistically insignificant for the sample of identical twins only. This corroborates the previous findings by Le et al. (2005) on the quantity dimension of human capital accumulation. They investigated the effects of the conduct disorder score on high school graduation and on years of education. For the third definition of conduct disorder we find conduct disorder to decrease the probability of high school graduation by 13 percent points using the sample of identical twins only. It is likely that this third measure of conduct disorder identifies individuals with severe behaviour problems

³ OLS-estimates for the sample of identical twins are very similar to the OLS-estimates for the total sample of twins. These results can be obtained from the authors on request.

which hinder their development. The difference in findings with Le et al. (2005) might be explained by the focus of this third measure on this specific group. The indicators 'grade repetition' and 'marks in high school' provide insight in the time needed for attaining a certain level and the performance in high school. The estimates on grade retention suggests that one additional point on the conduct disorder score increases the probability of grade retention with 1.1 percent point (0.9 for the sample of all twins). This estimate is remarkably close to the within sibling estimate of 0.8 percent point both for the US and Canada reported by Currie and Stabile (2007). Conduct disorder as classified by the APA definition increases the probability of grade retention with 9 percent points and with the third definition the estimated effect is 11 percent points. We also find a statistically significant effect of the conduct disorder score on marks in high school.

A cautionary note is in order for the results for the three indicators of positive human capital, because we do not have information on the exact timing of the dependent variables. It is not clear whether the conduct disorder behaviour, measured before the age of 18, predated or followed these outcomes. We address the issue of reverse causality in the section on the robustness checks.

To check for differences between men and women we re-estimated the main models with an interaction term of conduct disorder and gender. In line with Le et al. (2005) we find that the effect of conduct disorder on high school graduation is stronger for women than for men. However, for the third definition we find that for men the effect of conduct disorder on 'grade repetition' and on 'jail' is stronger than for women.

Table 4.1 Estimates of the effect of conduct disorder on positive and negative human capital

Conduct Disorder Score	Grade repetition	Marks high school	High school graduation	Attacking others	Arrested since 18	Jail
OLS	0.017 (0.003)*** [0.116]	- 0.049 (0.004)*** [-0.207]	- 0.026 (0.003)*** [-0.157]	0.041 (0.004)*** [0.266]	0.026 (0.003)*** [0.253]	0.014 (0.003)*** [0.253]
N	5286	5270	5288	2180	2178	2176
Fixed effect All twins	0.009 (0.003)*** [0.059]	- 0.038 (0.005)*** [-0.159]	- 0.013 (0.003)*** [-0.081]	0.037 (0.006)*** [0.241]	0.019 (0.004)*** [0.182]	0.011 (0.002)*** [0.193]
N	5286	5270	5288	2180	2178	2176
Fixed effect Identical twins	0.011 (0.005)** [0.077]	- 0.017 (0.009)** [-0.072]	- 0.010 (0.006) [-0.059]	0.029 (0.011)*** [0.189]	0.011 (0.007) [0.103]	0.007 (0.003)** [0.134]
N	2238	2238	2236	888	888	886
APA Definition						
OLS	0.096 (0.019)*** [0.089]	- 0.267 (0.026)*** [-0.155]	- 0.136 (0.019)*** [-0.113]	0.179 (0.022)*** [0.196]	0.124 (0.017)*** [0.201]	0.048 (0.011)*** [0.147]
N	5224	5210	5226	2140	2138	2136
Fixed effect All twins	0.057 (0.018)*** [0.053]	- 0.180 (0.031)*** [-0.105]	- 0.054 (0.020)*** [-0.045]	0.146 (0.029)*** [0.159]	0.076 (0.019)*** [0.122]	0.020 (0.010)** [0.060]
N	5224	5210	5226	2140	2138	2136
Fixed effect Identical twins	0.091 (0.026)*** [0.085]	- 0.025 (0.045) [-0.014]	- 0.036 (0.032) [-0.030]	0.162 (0.044)*** [0.177]	0.067 (0.028)** [0.108]	0.022 (0.013)* [0.067]
N	2220	2220	2218	876	876	874
Professional help						
OLS	0.201 (0.039)*** [0.090]	- 0.308 (0.049)*** [-0.086]	- 0.218 (0.040)*** [-0.087]	0.153 (0.045)*** [0.084]	0.125 (0.038)*** [0.103]	0.137 (0.035)*** [0.211]
N	5370	5354	5372	2246	2244	2242
Fixed effect All twins	0.105 (0.033)*** 0.047	- 0.157 (0.059)*** -0.044	- 0.126 (0.038)*** -0.050	0.022 (0.054) 0.012	0.061 (0.036)* 0.050	0.101 (0.017)*** 0.154
N	5370	5354	5372	2246	2244	2242
Fixed effect Identical twins	0.109 (0.048)** 0.049	- 0.088 (0.085) -0.025	- 0.130 (0.061)** -0.052	0.001 (0.085) 0.000	0.157 (0.054)*** 0.128	0.081 (0.025)*** 0.125
N	2274	2274	2272	916	916	914

Note: Each cell contains the estimate of a regression of a human capital measure on conduct disorder. All models control for birth weight, age, age squared, gender and parents' education. Standard errors in brackets, standardized coefficient in square brackets. ***/**/* significant at 1 %/5 %/10 %-level.

In general, we find the largest effects of conduct disorder for the third measure. The size of the estimates for the first and the second measure seems roughly comparable if we consider that

twins with conduct disorder according to the APA definition score approximately 5 points higher on the conduct disorder score. In addition, we constructed a variable similar to the APA definition based on the conduct disorder score only. This variable equaled 1 if the individual had a conduct disorder score of at least 3 and value zero if the individual had a lower conduct disorder score. Hence, the 12-month restriction has been dropped. We compared the estimation results of the effect of this new variable with the estimation results of the effect of the APA-definition. We find that the size of the estimates is somewhat smaller when using this new variable. In particular, the size of the coefficient estimates increases and becomes similar to the results with the APA-definition when we recode individuals with conduct disorder scores of 3 (or 3 and 4) from one to zero.

Non linearity

We also investigated whether the effects are non linear by estimating the effects of dummies for having a conduct disorder score of more than 5 or at least 10.⁴ This corresponds to the 90th and 99th percentile. It should be noted that focusing on these top percentiles reduces the variation in conduct disorder that can be used, especially in the sample of identical twins. For most outcomes the estimates suggest that the effects of conduct disorder are quite linear. For instance, for the sample of all twins the within estimates using the dummy for at least 10 conduct disorder behaviours (above the 98 percentile) seem quite comparable with most estimates in table 4.1. Hence, there is no clear evidence for non linear effects of conduct disorder.

Effects of four subscales of conduct disorder

The APA criteria suggest that conduct disorder has four underlying components: aggression to people and animals, destruction of property, deceitfulness or theft and serious violations of rules. We investigated the effects of these components by decomposing the conduct disorder score into these four groups of the APA criteria. The estimation results suggest that ‘serious violations of rules’ are the most important component of conduct disorder for the accumulation of positive human capital. This is not very surprising as this subscale includes items like ‘misbehaved’, ‘wagged school’ and ‘suspended/expelled’. For the other subscales the estimates do not show a clear pattern. The estimates of the effect on negative human capital suggest that all subscales are important. The size of the effects of ‘aggression’ and ‘destruction’ on ‘attacking others’ and ‘failing to pay debts’ is remarkable. ‘Destruction’, ‘deceit/theft’ and ‘violation of rules’ seem the most important factors for the probability of arrest or jail.

The effect of the timing of conduct disorder behaviours

⁴ The estimation results on non linearity and the four subscales of conduct disorder can be obtained from the authors on request.

Several recent studies suggest that the timing of intervention programmes for disadvantaged children is important. For instance, Carneiro and Heckman (2003) show that socio-economic differences in cognitive and non-cognitive abilities appear early in life and widen over the life cycle of the child. Currie and Stabile (2007) find that early mental health problems have large significant effects on cognitive test scores even controlling for later mental health problems. Loeber (1982) suggests that the earlier a child began displaying antisocial behaviour, the more likely he or she is to persist in such behaviour. Farmer (1995), on the other hand, does not support the hypothesis that earlier onset is associated with more deleterious outcomes. She finds that later onset, rather than earlier, is associated with earlier school leaving and lower level qualifications.

We estimated the effect of the timing of conduct disorder on human capital by including three dummies for the age-of-onset ranges: 3-12, 13-14 and 15-18 (see equation 3.2). The estimates in table 4.2 suggest that the timing of conduct disorder behaviours matters for human capital accumulation, especially for negative human capital. An earlier occurrence of conduct disorder seems more detrimental. The largest estimates of the effect of conduct disorder on 'attacking others' are obtained for the two youngest age ranges. In addition, an early age of onset (≤ 12 year) has a strong effect on the probability of incarceration, whereas a later age of onset has no effect on this outcome. The timing of conduct disorder on positive human capital seems less important. The youngest age range seems most important for high school graduation but the estimates for identical twins are statistically insignificant. We also estimated models in which we included the years of conduct disorder before the age of 18 and the square of these years. These estimates confirm that the effect of conduct disorder on human capital is more deleterious if these behaviours occur earlier but the impact of the timing differs between outcomes (see appendix A.2).

We also investigated whether there is a relationship between the socioeconomic background of the twins and the effect of the timing of conduct disorder on human capital. As wealthier families have more resources to mitigate conduct disorder problems, we might expect smaller effects for twins from these families. We separated our sample in families with a father who completed high school and families with a father who did not complete high school. However, estimates which include an interaction between socioeconomic background and the timing of conduct disorder show little effect of socioeconomic background. We only find that conduct disorder has a larger effect on the probability of physically attacking others in families in which the father did not complete high school than in other families. Currie and Stabile (2007) also find little evidence that parental income mitigates negative effects of mental health problems of children.

The findings in this section are based on self-reports. We are not aware of evidence regarding the accuracy of self-reported data on the timing of conduct disorder. Therefore, a cautionary note on measurement error is in order.

Table 4.2 The effect of the timing of conduct disorder on positive and negative human capital

APA Definition	Grade repetition	Marks high school	High school graduation	Attacking others	Arrested since 18	Jail
OLS						
Age of onset						
≤ 12 year	0.052 (0.034)	-0.255 (0.057)***	-0.180 (0.037)***	0.246 (0.043)***	0.133 (0.036)***	0.084 (0.029)***
13-14 year	0.161 (0.033)***	-0.329 (0.048)***	-0.211 (0.033)***	0.176 (0.038)***	0.130 (0.030)***	0.032 (0.016)**
15-17 year	0.073 (0.025)***	-0.233 (0.033)***	-0.065 (0.025)***	0.150 (0.031)***	0.118 (0.024)***	0.040 (0.014)***
N	5224	5208	5224	2138	2136	2134
Fixed effect All twins						
≤ 12 year	0.049 (0.033)	-0.154 (0.059)***	-0.052 (0.038)	0.231 (0.052)***	0.054 (0.035)	0.044 (0.017)***
13-14 year	0.084 (0.028)***	-0.206 (0.050)***	-0.121 (0.033)***	0.179 (0.047)***	0.097 (0.031)***	0.004 (0.015)
15-17 year	0.043 (0.024)*	-0.173 (0.043)***	-0.012 (0.028)	0.095 (0.038)**	0.074 (0.025)***	0.017 (0.012)
N	5224	5208	5224	2138	2136	2134
Fixed effect Identical twins						
≤ 12 year	0.069 (0.050)	-0.114 (0.087)	-0.031 (0.062)	0.225 (0.083)***	0.067 (0.053)	0.100 (0.024)***
13-14 year	0.087 (0.042)**	0.003 (0.073)	-0.081 (0.052)	0.295 (0.074)***	0.079 (0.047)*	-0.008 (0.021)
15-17 year	0.103 (0.034)***	-0.005 (0.059)	-0.013 (0.042)	0.074 (0.055)	0.061 (0.035)*	0.005 (0.016)
N	2220	2220	2218	876	876	874

Notes: Estimates of a separate regression of a human capital measures on three dummies for the age of onset of conduct disorder using the same controls as in table 5.1. Standard errors in brackets. ***/**/* significant at 1 %/5 %/10 %-level.

5 Robustness checks

Reverse causality

A caveat with respect to the findings on positive human capital in table 5.1 is that we do not observe the exact timing of the dependent variables. Therefore, we cannot be sure whether the conduct disorder behaviours predated or followed these outcomes. However, with respect to the APA-definition of conduct disorder respondents were asked: ‘Did these behaviours cause problems for you at school?’. This question enables us to address the issue of reverse causality, because it seems likely that for respondents who answered ‘yes’ to this question the causation runs from conduct disorder to human capital accumulation. We therefore re-estimated the models for positive human capital after excluding twins with conduct disorder according to the APA-definition who answered ‘no’ on this question. Hence, we excluded twins for which the causation of the effects is not clear. The results are shown in table 5.1.

Table 5.1 The effect of conduct disorder (APA Definition) on positive human capital after excluding twins that state that their conduct disorder behaviours did not cause problems at school

APA Definition	Grade retention	Marks high school	High school graduation
OLS	0.156 (0.035)***	-0.375 (0.048)***	-0.207 (0.036)***
N	4402	4388	4406
FE All	0.102 (0.030)***	-0.276 (0.054)***	-0.092 (0.035)***
N	4402	4388	4406
FE Identical	0.098 (0.043)**	-0.139 (0.079)*	-0.071 (0.055)
N	1928	1928	1928

We observe that the pattern of findings in table 5.1 is similar to the findings in table 4.1, although the size of most estimates is larger. Hence, our findings are robust for excluding respondents for which the causation of the effects is not clear.

Missing values due to the routing of the questionnaire

In section 3 we noted that due to the routing of the questionnaire twins with a conduct disorder score of zero did not answer the question on negative human capital. These missing values may bias the estimates downward as it seems likely that twins with no childhood conduct disorder behaviour will on average accumulate more positive and less negative human capital than twins with problem behaviour early in life. We checked the sensitivity of the results by imputing the sample mean of the specific outcome for twins with missing values and a conduct disorder score of zero. Table 5.2 shows the estimation results for the APA definition of conduct disorder.

Table 5.2 Estimates of the effect of conduct disorder (APA definition) on positive and negative human capital after imputation of missing values due to the routing of the questionnaire

APA definition	Attacking others	Arrested since 18	Jail
OLS	0.176 (0.019)***	0.115 (0.015)***	0.043 (0.010)***
N	5184	5180	5178
Fixed effect All twins	0.158 (0.019)***	0.084 (0.012)***	0.025 (0.006)***
N	5184	5180	5178
Fixed effect Identical twins	0.166 (0.030)***	0.060 (0.018)***	0.021 (0.008)**
N	2200	2198	2196

Note: Estimates from regressions of negative human capital on the APA definition of conduct disorder. Missing values for human capital are imputed with mean values.

After the imputation of these missing values, all estimates become statistically significant. We also observe that the size of most estimates increases. This suggests that due to the routing of the questionnaire we underestimate the deleterious effects of conduct disorder on negative human capital. Hence, the size of the estimates in the previous section may be considered as a lower bound of the true effects of conduct disorder on these outcomes.

Excluding pairs of twins with large differences in educational attainment

As a third robustness check we excluded pairs of twins with large differences in educational attainment (more than 5.5 years difference within pairs). These differences might indicate that these twins are quantitatively different from the rest of the sample and introduce heterogeneity which will confound the effects we are looking for. We find that the estimates hardly change after the exclusions of these pairs of twins.

Comorbidities

Many children with a conduct disorder may have coexisting conditions such as mood disorders (depression), anxiety, ADHD or substance use disorders. These conditions might bias our estimates. To our knowledge, only Currie and Stabile (2007) have addressed the issue of comorbidity so far. They find that the incidence of hyperactivity, depression and conduct disorder are correlated across individuals and, therefore, include all mental health scores plus the total combined score in their estimation model. This yields estimation results that are consistent with the results from models that only include one mental health problem, although the estimates are less precise. Their results on the effect of conduct disorder in the United States become statistically insignificant after including all other mental health scores.

To address the issue of comorbidity, we included controls for a large set of self-reported physical and mental health conditions before the age of 14 measured in the first survey. Respondents were asked whether they had had any of the following before they were 14 years old: Migraine/sick headaches; Depression; Asthma; Alcohol problem; Sudden, involuntary

movements (tics); Sudden, involuntary noises or utterances (vocal tics); Involuntary utterance of swear words; Autism; Stuttering or stammering; Diabetes; Epilepsy or suspected epilepsy; Seizure, convulsion; Treatment for schizophrenia; Treatment for manic-depression; Cleft lip; Melanoma; Moles removed. Unfortunately, our data do not include measures of hyperactivity. We included dummies for these 17 conditions in the models that use the APA definition of conduct disorder. The estimation sample becomes smaller due to missing values on these dummies. A comparison of these estimates with estimates of the main model using this smaller sample shows that the estimated effect of conduct disorder on positive and negative human capital hardly changes after the inclusion of these 17 dummies.⁵ In addition, we are able to use measures of depression from the second survey. In particular, respondents were asked: Has there ever been two weeks or more when you were depressed or down most of the day, nearly every day?; Has there ever been two weeks or more when you were a lot less interested in most things or unable to enjoy the things you used to enjoy, most of the day nearly every day? Moreover, the age of onset was asked. From these variables we constructed a dummy variable for depression before the age of 18. Table 5.3 shows the estimation results of the models that include conduct disorder and an indicator of depression before the age of 18.

⁵ Results can be obtained from the authors on request.

Table 5.3 Estimates of the effect of conduct disorder (APA Definition) on positive and negative human capital after including an indicator of depression before the age of 18

APA Definition	Grade retention	Marks high school	High school	Attacking others	Arrested since 18	Jail
OLS						
Conduct disorder	0.095 (0.019)***	-0.266 (0.026)***	-0.135 (0.019)***	0.176 (0.022)***	0.122 (0.017)***	0.045 (0.010)***
Depression	0.023 (0.018)	-0.016 (0.032)	-0.017 (0.019)	0.050 (0.028)*	0.046 (0.021)**	0.044 (0.015)***
N	5224	5210	5226	2140	2138	2136
Fixed effect All twins						
Conduct disorder	0.055 (0.018)***	-0.179 (0.031)***	-0.052 (0.020)**	0.143 (0.029)***	0.073 (0.020)***	0.018 (0.009)*
Depression	0.029 (0.019)	-0.021 (0.034)	-0.040 (0.022)*	0.052 (0.039)	0.049 (0.026)*	0.047 (0.013)***
N	5224	5210	5226	2140	2138	2136
Fixed Effect Identical twins						
Conduct disorder	0.087 (0.026)***	-0.019 (0.045)	-0.030 (0.032)	0.155 (0.044)***	0.068 (0.028)**	0.021 (0.013)
Depression	0.062 (0.027)**	-0.093 (0.048)*	-0.100 (0.034)***	0.133 (0.063)**	-0.022 (0.040)	0.013 (0.019)
N	2220	2220	2218	876	876	874

Note: Estimates from regressions of human capital on the APA definition of conduct disorder.

Controlling for depression before the age of 18 does not change the results. The estimated effects of conduct disorder in table 5.3 are slightly smaller but very similar to the previous results in table 4.1. Depression during childhood also seems important for both positive and negative human capital, which confirms findings by Fletcher (2008). Remarkably, for the indicators of positive human capital the largest effects of depression are found within pairs of identical twins. These findings suggest that the estimates of the effect of conduct disorder on human capital are robust for including measures of coexisting physical or mental health conditions. However, some caution is needed because there might be many coexisting conditions which are typically difficult to observe. In addition, we do not have measures of hyperactivity which has been found to be comorbid with conduct disorder.

6 Conclusions

In this paper, we examine the relationship between conduct disorder problems before the age of 18 and human capital accumulation. The estimates suggest that conduct disorder has a strong effect on violent and criminal behaviour since the age of 18. In addition, we find that conduct disorder has a negative effect on positive human capital. For instance, depending on the measure being used, we find that conduct disorder reduces high school graduation with 4 to 13 %-points and increased the probability of being arrested with 7 to 11 %-points. Our data also provide the opportunity to investigate whether the timing of conduct disorder matters. In general, we find that earlier occurrence of conduct disorder is more deleterious. Various robustness checks suggest that these estimates may be lower bounds of the true effect of conduct disorder on human capital. For instance, due to the routing of our survey twins with a conduct disorder score of zero did not answer most questions on negative human capital. The size of the estimates increases if we impute the sample means of negative human capital outcomes for these twins.

Previous research (Currie & Stabile, 2007) indicates that conduct disorder before the age of 18 can have large deleterious longer-term effects on positive and negative human capital. Our findings, which are based on a much richer set of indicators of conduct disorder and on data of twins instead of data of siblings, corroborate these results. The magnitude of their estimates on grade repetition is very similar to our findings and those for young adult delinquency are in the same range as our estimates on attacking others and probability of being arrested⁶.

Le et al. (2005), who used the same Australian data set, also investigated the effect of the conduct disorder score on the quantity of human capital accumulation (high school graduation and years of education). For the sample of identical twins they find that the estimated effect is not statistically significant. We find basically the same effect for this measure. However, by investigating a broader set of indicators of human capital and by using three measures of conduct disorder we find evidence for detrimental effects of conduct disorder, even in the sample of identical twins. These findings lead us to conclude that conduct disorder decreases investment in positive human capital and increases negative human capital.

An important issue that is still unresolved in the literature relates to the source of the variation in conduct disorder, in particular the variation within pairs of twins. The literature is not clear about the origin of conduct disorder and mentions a range of factors related to prenatal and birth conditions and the environment for child rearing (e.g. Loeber et al. 1998). It is difficult to assess to which extent these variables might also vary within pairs of twins and, accordingly, generate variation in conduct disorder within pairs of twins. With respect to the nurture-nature debate the data at hand might give some insight. We find that approximately two thirds of the variation in conduct disorder is captured by fixed effects. For the outcome

⁶ Their estimates using a 16 point scale range between 0.01 and 0.034, where our estimates using a 21 point scale for identical twins range from 0.011 to 0.029

variables we find that fixed factors capture more than 50% of the variation in positive human capital and between 10 to 30% of the variation in negative human capital.⁷ The variation captured by fixed effects is informative on the maximum share that can be attributed to nature. The remaining part is apparently still ‘up for grab’ after childhood. This suggests that, especially with respect to negative human capital, there is scope for environmental factors or programs to improve outcomes.

The empirical findings in this paper lead us to conclude that early behavior problems have detrimental longer term effects on human capital, especially on violent and criminal behavior later in life.

⁷ The intra class correlation for identical twins are 0.49 for high school graduation, 0.51 for marks in high school, 0.60 for grade repetition, 0.10 for attacking others, 0.18 for arrest and 0.29 for jail.

References

Ashenfelter, O. and A.B. Krueger, 1994, Estimates of the economic return to schooling from a new sample of twins, *American Economic Review*, 84(5), 1157-73.

Australian Centre for Community Child Health, 2006, Royal Children's Hospital Melbourne, Research Centre of the Murdoch Childrens Research Institute, Media Information.

Baker, L.A., Treloar, S.A., Reynolds, C.A., Heath, A.C., Martin, N.G., 1996, Genetics of educational attainment in Australian Twins: Sex differences and secular changes, *Behavior Genetics*, 26(2), 89-102.

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

Borghans, L., A.L. Duckworth, J.J. Heckman and B. ter Weel, 2008, The economics and psychology of personality traits, *Journal of Human Resources*, vol. 43, no. 4, pp. 972-1059.

Bucholz, K.K., Cadoret, R., Cloninger, C. R., Dinwiddle, S.H., Hesselbrock, V.M., Nurnberger, J.L., Jr. Reich, T., Schmidt, I., Schuckit, M.A., 1994, A new semi-structured psychiatric interview for use in genetic linkage studies: A report on the reliability of the SSAGA, *Journal of Studies on Alcohol*, 55, 149-158.

Currie, J. and M. Stabile, 2006, Child Mental Health and Human Capital Accumulation: The Case of ADHD, *Journal of Health Economics*, Vol. 25, No. 6, pp. 1094-1118.

Currie, J. and M. Stabile, 2009, Mental Health in Childhood and Human Capital, in: *The Problems of Disadvantaged Youth: An Economic Perspective*, University of Chicago Press.

Farmer, E.M.Z., 1995, Extremity of Externalizing Behaviour and Young Adult Outcomes, *Journal of Child Psychology*, Vol. 36, No. 4, pp. 617-632.

Farrington, D. (1973). Self-reports of deviant behavior: Predictive and stable?, *Journal of Criminal Law and Criminology*, 64(1): 99-110.

Fletcher, J. and B.L. Wolfe, 2007, Long-term consequences of childhood ADHD on criminal activities, University of Wisconsin-Madison.

Fletcher, J. and B.L. Wolfe, 2008, Child Mental Health and Human Capital Accumulation: The Case of ADHD Revisited, *Journal of Health Economics*, 27, pp. 794–800.

Griliches, Z., 1979, Sibling models and data in economics: beginnings of a survey, *Journal of Political Economy*, 87(5), pp. S37-S64.

Hardt, R. and S. Hardt, 1997, On determining the quality of the delinquency self-report method, *Journal of Research in Crime and Delinquency*, July: 247-257.

Heath, A.C., Madden, P.A.F. and N.G. Martin, 1998, Assessing the effects of cooperation bias and attrition in behavioural genetic research using data-weighting, *Behavior Genetics*, 28(6), 415-427.

Hesselbrock, M., Easton, C., Bucholz, K.K., Schuckit, M., Hesselbrock, V., 1999, A validity study of the SSAGA, a comparison with SCAN, *Addiction*, 94, 1361-1370.

Horney, J. and I. Marshall, 1992, An experimental comparison of two self-report methods for measuring lambda, *Journal of Research in Crime and Delinquency*, 29(1): 102-121.

Huizinga, D. and D.S. Elliott, 1986, Reassessing the reliability and validity of self-report delinquency measures, *Journal of Quantitative Criminology*, 2, 293-327.

Le, A.T., P.W. Miller, A.C. Heath and N. Martin, 2005, Early childhood behaviours, schooling, and labour market outcomes: estimates from a sample of twins, *Economics of Education Review*, Vol. 24, pp. 1-17.

Loeber, R., 1982, The Stability of Antisocial and Delinquent Child Behaviour: A Review, *Child Development*, Vol. 53, pp. 1431-1446.

Loeber, R., Farrington, D.P., Stouthamer-Loeber, M., & Van Kammen, W.B, 1998, Antisocial behavior and mental health problems: Explanatory factors in childhood and adolescence. Mahwah, NJ: Lawrence Erlbaum Associates.

Maddux J. and D. Desmond, 1975, Reliability and validity of information from chronic heroin users, *Journal of Psychiatric Research*, 12(2): 87-95.

Mieczkowski, T., 1990, The accuracy of self-reported drug use: An evaluation and analysis of new data, in R. Weisheit (Ed.), *Drugs, crime and the criminal justice system*, pp. 275-302. Cincinnati: Anderson.

Miller, P., C. Mulvey and N. Martin, 1995, What Do Twins Studies Reveal about the Economic Return to Education? A Comparison of Australian and U.S. Findings, *American Economics Review*, Vol. 85, No. 3, pp. 586-599.

Miller, P., C. Mulvey and N. Martin, 2006, The return to schooling: Estimates from a sample of young Australian twins, *Labour Economics*, 13, pp. 571-587.

Searight, H. R., Rottnek, F., Abby, S.L. , 2001, Conduct disorder: diagnoses and treatment in primary care, *American Family Physician*, 63 (8), 1579-1588.

Slutske, W.S., Heath, A.C., Dinwiddie, S.H., Madden, P.A.F., Bucholz, K.K., Dunne, M.P., Statham, D.J. and N.G. Martin, 1997, Modelling genetic and environmental influences in the etiology of conduct disorder: a study of 2,682 adult twin pairs, *Journal of Abnormal Psychology*, vol 106 (2), 266-279.

Taylor, N., 2006, Juveniles in Detention in Australia, 1981-2005, Technical and Background Paper 22, Australian Institute of Criminology.

Thornberry, T.P. and M.D. Krohn, 2000, The self-report method for measuring delinquency and crime, *Criminal Justice 2000*, Vol. 4, pp. 33-83.

Webbink, H.D., P.W.C. Koning, S. Vujić and N. Martin, 2008, Why are criminals less educated than non-criminals?, CPB Discussion Paper 114.

Weiss, R., L. Najavits, S. Greenfield, J. Soto, S. Shaw and D. Wyner, 1998, Validity of substance use self-reports in dually diagnosed outpatients, *American Journal of Psychiatry*, 155, 127-128.

Appendix

A.1 Data collection and external validity

In 1980-1982 a sample of 4,262 twin pairs, born between 1964 and 1971, were registered with the ATR as children by their parents in response to media appeals and systematic appeals through the school system. In 1989-1992, when the twins were 18-25 years old, the first survey by mailed questionnaire was conducted, called Alcohol Cohort 2. The response rate of this questionnaire survey was 63%. In 1996-2000, the second survey was launched, called TWIN89. Telephone interviews were completed with 6,267 individuals, 2,805 men (889 complete and 1,027 incomplete pairs) and 3,462 women (1,215 complete and 1,032 incomplete pairs), who were 30 years old on average (range from 24 to 39) at the time of the interview. The individual response rate for this telephone interview was 86%. The interview was based upon a modified version of the SSAGA (Semi-structured Assessment of the Genetics of Alcoholism). Modifications were made to the SSAGA to incorporate DSM-IV criteria as well as to adapt it for telephone use. Interviews were administered by trained lay interviewers who were blind to the psychiatric status of the co-twin. Interviews were supervised by a qualified clinical psychologist with 4 years of experience who reviewed all interview protocols. In addition, all interviews were tape-recorded and random interview tapes were reviewed for quality control.

Le et al. (2005) compared this twin sample with data of Australian individuals from the *Youth in Transition* surveys and concluded that the patterns in relation to school leaving decisions are broadly similar to those obtained from samples representative of the general population. Slutsky et al. (1997) noted that twins seem representative of the general population with respect to conduct disorder. They also compared the rates of conduct disorder in the so-called older cohort of Australian twins (we use the younger cohort) with the self-reported rate of conduct disorder for the general population of the US measured in the National Comorbidity Study⁸. They concluded that the Australian sample did not consistently differ with rates estimated for the general population of the United States.

Another issue is that severe cases of conduct disorder might be under-sampled or have a higher probability of attrition. We compared the rates of conduct disorder among twins concordant for participation in the second survey with the rates of conduct disorder among twins whose co-twin did not participate in the second survey. If twins with conduct disorder were systematically under-sampled, higher rates of conduct disorder would be expected among twins whose co-twin did not participate in the survey (assuming that conduct disorder status is correlated in twin pairs). We found however no difference between the two groups of twins in our sample. Slutsky et al. (1997) and Heath et al. (1998) also found no evidence for cooperation bias or attrition bias in the so-called older cohort of Australian twins.

⁸ We do not include the data of the older cohort in this paper because questions on negative human capital have not been asked.

A.2 The measurement of conduct disorder

Table A.1 shows the criteria for the diagnosis of conduct disorder such as defined by The American Psychiatric Association (APA).

Table A.1 DSM-IV Diagnostic Criteria for Conduct Disorder^a

A Repetitive and persistent pattern of behaviour in which the basic rights of others or major age-appropriate societal norms or rules are violated, as manifested by the presence of three (or more) of the following criteria in the past 12 months, with at least one criterion present in the past 6 months:

Aggression to people and animals

1. Often bullies, threatens, or intimidates others
2. Often initiates physical fights
3. Has used a weapon that can cause serious physical harm to others (e.g., a bat, brick, broken bottle, knife, gun)
4. Has been physically cruel to people
5. Has been physically cruel to animals
6. Has stolen while confronting a victim (e.g., mugging, purse snatching, extortion, armed robbery)
7. Has forced someone into sexual activity

Destruction of property

1. Has deliberately engaged in fire setting with the intention of causing serious damage
2. Has deliberately destroyed other's property (other than by fire setting)

Deceitfulness or theft

1. Has broken into someone else's house, building, or car
2. Often lies to obtain goods or favours to avoid obligations (i.e., "cons" others)
3. Has stolen items of nontrivial value without confronting a victim (e.g., shoplifting, but without breaking and entering; forgery)

Serious violations of rules

1. Often stays out at night despite parental prohibitions, beginning before age 13 years
2. Has run away from home overnight at least twice while living in parental or parental surrogate home (or once without returning for a lengthy period)
3. Is often truant from school, beginning before age 13 years.

B The disturbance in behaviour causes clinically significant impairment in social, academic, or occupational functioning.

C If the individual is age 18 years or older, criteria are not met for antisocial personality disorder.

Specify severity:

Mild: few if any conduct problems in excess of those required to make the diagnosis and conduct problems cause only minor harms to others.

Moderate: number of conduct problems and effect on others intermediate between "mild" and "severe."

Severe: many conduct problems in excess of those required to make the diagnosis or conduct problems cause considerable harm to others.

^a Source: American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed. Washington, DC: American Psychiatric Association, 1994:98-99.

Self-reported information on 21 statements that reflect behavioural problems before the age of 18 has been collected in the second survey when the twins were on average 30 years old. We created our first measure of conduct disorder by summing occurrences of these 21 statements⁹. Our second measure of conduct disorder is based on the question ‘Did you do at least 3 of these things within the same 12-month period?’. Twins who responded ‘yes’ were coded as 1, twins who responded no or twins with a conduct disorder score of 0, 1 or 2 were coded as 0. As this measure perfectly matches the APA definition of conduct disorder we will use it as our main measure in the analysis. We call this measure the ‘APA Definition’. Our survey also asked ‘How old were you the first time you did at least 3 of these things within the same 12-month period?’. We use this information to investigate the effect of the timing of conduct disorder. Our third measure of conduct disorder is based on the question ‘Did you ever see a doctor, psychologist, social worker or any other professional like that because of these behaviours?’. Twins who responded ‘yes’ were coded as 1, twins who responded ‘no’ or twins who were coded as 0 on the APA definition were coded as 0 on this third measure. We define this measure as ‘Professional help’. Seeking professional help may signal that conduct disorders hinder normal development and human capital accumulation. Obviously, this measure is endogenous as people choose or are advised to seek professional help.

The second measure of conduct disorder, which is directly based on the APA definition, might be a better indicator than the conduct disorder score because of the additional restriction on the time range for the occurrence of the conduct disorders. Small variations on the conduct disorder score (the first definition) might not reflect real differences in conduct disorder if the conduct disorder behaviors occur with relatively large time lags. The psychiatric handbook (DSM-IV) emphasizes that at least three such behaviors should occur within the same 12-month period to make the diagnosis of conduct disorder. The advantage of the third measure is that seeking professional help is a clear signal that the behavior of the child is considered problematic, making it likely that this indicator really measures a difference in conduct disorder.

⁹ We use the same 19 statements as in Le et al. (2005) and also included ‘misbehaved’ and ‘suspended/expelled’.

Table A.2 The effect of the timing of conduct disorder on human capital

APA Definition	Grade repetition	Marks high school	High school graduation	Attacking others	Arrested since 18	Jail
OLS						
Years of CD	0.040 (0.009)***	- 0.110 (0.013)***	- 0.058 (0.009)***	0.061 (0.011)***	0.042 (0.009)***	0.012 (0.007)*
Years of CD squared	- 0.003 (0.001)***	0.009 (0.002)***	0.004 (0.001)***	- 0.004 (0.002)**	- 0.003 (0.001)**	- 0.000 (0.001)
N	5222	5208	5224	2138	2136	2134
Fixed effect All						
Years of CD	0.024 (0.009)***	- 0.083 (0.016)***	- 0.036 (0.010)***	0.060 (0.014)***	0.035 (0.009)***	0.011 (0.005)**
Years of CD squared	- 0.002 (0.001)*	0.008 (0.002)***	0.004 (0.001)***	- 0.005 (0.002)**	- 0.003 (0.001)***	- 0.001 (0.001)**
N	5222	5208	5224	2138	2136	2134
Fixed effect Identical						
Years of CD	0.044 (0.013)***	- 0.000 (0.024)	- 0.033 (0.017)*	0.056 (0.023)**	0.032 (0.014)**	0.004 (0.007)
Years of CD squared	- 0.004 (0.002)**	- 0.001 (0.003)	0.005 (0.002)*	- 0.003 (0.003)	- 0.003 (0.002)*	0.000 (0.001)
N	2220	2220	2218	876	876	874

Notes: Estimates of separate regression of a human capital measures on years of conduct disorder and years of conduct disorder squared using the same controls as in table 5.1. Standard errors in brackets.

***/**/* significant at 1 %/5 %/10 %-level.