COGNITION, CANNABIS AND WAGES *

Gabriella Conti
Department of Economics
University of Chicago

First version: August 2007
This version: January 2010

PRELIMINARY DRAFT
PLEASE DO NOT CITE WITHOUT PERMISSION

Abstract

This paper explores the cognitive origins of cannabis use, and revisits the effect of cannabis use on wages. It provides evidence of a positive and significant relationship between cognitive ability and cannabis use and confirms the wage returns to cognitive ability. It then shows that the relationship between cannabis use and wages observed in cross-sectional estimates is spurious and derives instead from the omission of cognitive ability.

JEL code: I12, J24
Keywords: Cannabis, cognitive ability, wage, factor models.

*Gabriella Conti, Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, Illinois 60637, `gconti@uchicago.edu`. This work was supported by the Economic and Social Research Council, award no. PTA030200300812. I thank Wiji Arulampalam, Alison Booth, James Heckman and Steve Pudney for comments. I thank Brian Dodgeon for help in constructing the test scores data.

1 Introduction

"I did break the law... I was wrong... drugs are wrong."

On 19th July, 2007, the headline news in all British newspapers was a declaration by the former Home Office Secretary Jacqui Smith. She admitted to having smoked cannabis "a few times" while she was a student at Oxford University in the 1980s. Her public disclosure came the day after the Prime Minister Gordon Brown announced a review of the UK drugs strategy, and foreshadowed the announcement of a new government strategy against crime. However, Jacqui Smith is not the only successful person who has tried cannabis in her adolescent years: a number of politicians have made the same admission. Now the question is: should we be worried about this? Or should this just tell us that not all drug users are alike, and that having experimented with drugs does not hinder success in life?

The conventional wisdom holds that use of illicit substances adversely affects labour market productivity (see Culyer [1973], for one of the pioneering studies). Although this seems a priori reasonable, the empirical literature has yet to establish a consensus on the topic. This paper revisits the relationship between cannabis use and wages (as proxy for labour market productivity) using a rich longitudinal data source for the United Kingdom and looking at a previously unexplored factor - cognitive ability - as linking the processes of drug consumption and wage determination.[1] Drug policy is hotly debated in Great Britain nowadays: since the depenalisation of cannabis in 2004, there has recently been a reclassification from class C to class B, despite opposite advice given by the Advisory Council on the Misuse of Drugs. This study provides the first systematic analysis of the relationship between cannabis use and wages using a dataset still unexploited for this purpose (the British Cohort Study) and newly developed factor models for studying latent cognitive traits (see Heckman et al. [2006]).

The rest of the paper is organised as follows. In the next section we discuss the literature on drug use, cognitive ability and labour market success. In section 3 we describe the data used in this paper, and we provide some first descriptive evidence on the patterns of drug use and cognitive ability. We provide preliminary OLS-based evidence in section 4, and we outline the econometric methodology that we use in section 5, leaving the discussion of the estimates to section 6. Section 7 concludes.

2 Previous Literature

This paper combines two strands of literature: a literature on the determinants of drug use, and its effect on human capital and wages, and a literature on the determinants and the effects of cognitive ability.

[1] Becker and Murphy [1988], in their seminal paper, suggest that there might be unobserved factors affecting both drug use and wages. However, up to this point, the empirical literature has been unable to shed some light on this issue.

2.1 Drug and cognitive ability

All the studies which have looked at the relationship between cognitive ability and cannabis use are based on US data. Unfortunately, in most of them drug use and cognitive ability are measured at points very close in time, posing a serious threat to credible interpretations of the relationship between them in terms of causality. The direction of the association is not clear: Sander [1998], for example, finds no association between mathematics test scores and marijuana use at high school, using the High School and Beyond data, while DeSimone [2002] finds a positive effect of the AFQT measured in 1980 on marijuana use in 1984, using the National Longitudinal Survey of Youth data.[2] In studies which attempt to address causality issues, the cross-sectional association between marijuana use and cognitive ability is attenuated. Heckman et al. [2006], for example, use factor models to control for endogeneity and measurement error in the test scores, and find that the cognitive factor (as measured by test scores administered in 1979 and 1980, when the sample members were aged 14 to 23) is not a strong predictor of marijuana use in 1979.

The effect of cannabis on cognitive functioning, instead, has been mainly studied in the medical literature, which has not achieved a definite consensus: Solowij [1998] attributes the diversity of findings in the literature to differences in the measurement of marijuana consumption, and the length of time for which is taken. In the economic literature, Pacula et al. [2003] are among the few researchers who try to address the issue of unobserved heterogeneity. They employ a difference-in-differences specification which relates changes in student achievement to changes in marijuana use over the same time period: however, there remains a major shortcoming. Even if the difference-in-differences approach takes care of unobserved time-invariant factors affecting both drug use and achievement, it does not rule out the possibility that endogeneity might be a serious issue also when looking at changes rather than at levels.[3]

There is no greater consensus in the literature which looks at the relationship between marijuana use and schooling outcomes: the results of previous studies are mixed, and some research has been suggestive of a positive effect of attending college on drug access (Chen [2006]), while others (McCaffrey *et al.* [2008]) have shown a positive association between heavy and persistent marijuana use and high school dropout.[4] Conti *et al.* [2010] show that, while more educated individuals are more likely to have experimented with cannabis use, post-compulsory education has a negative causal effect on lifetime prevalence.

While we try to shed some light on the relationship between cognitive ability and cannabis use exploiting the richness and the time span of our data, we show that two aspects are crucial: the age at which the former is measured, and the way in which cannabis consumption is measured - in particular, whether frequency of use is taken into account.[5]

[2]He labels this effect as "unexpected" and conjectures it captures "an income effect from socioeconomic factor". Here we show that the effect is still present when controlling for socioeconomic background. In addition, while a positive association between cognitive ability and drug use comes as a surprise in the economic literature, it has already been recognised for a long time in the psychiatric literature (see Kolb [1925] and recently Heyman [2009]).

[3]It is useful to notice that in the NELS:88 (National Education Longitudinal Study) were administered tests in reading, mathematics, history and science that, for both the timing and the content of the tests themselves, reflect more academic learning than innate cognitive abilities.

[4]Indeed, the literature seems to be suggestive of both a positive link between early involvement with marijuana and high school dropout (Mensch and Kandel [1988]) and the opposite (Sander [1998]).

[5]Our dataset is unique as it covers a very long time period and it has a very rich set of background

2.2 Drug and labour market outcomes

Cognitive ability is also an important factor in the relationship between drug use and labour market outcomes.[6] Strikingly, most of the US literature on the topic points to a positive cross-sectional relationship between drug use and wages (Gill and Michaels [1991b], Kaestner [1991], Register and Williams [1992]), which tends to disappear when unobserved heterogeneity is properly taken into account. Kaestner [1994], indeed, is the first to control for unobserved heterogeneity in the drug use-earnings relationship, using newly-released data from the 1984 and the 1988 waves of the NLSY: the positive relationship between drug use and wages present in the cross-sectional estimates disappears when fixed-effects methods are used. Burgess and Propper [1998], also using the NLSY, find that soft drug use during adolescence has basically no effect on wages of men in their twenties or thirties. For the UK, MacDonald and Pudney [2000] also find very limited evidence to sustain any relationship between drug use and occupational attainment (their measure of labour market productivity); in a closely related paper (MacDonald and Pudney [2001]), they find a very weak positive association between 'soft' drug use and occupational success. Van Ours [2007], instead, is one of the few studies to find a negative and significant association between drug use and wages, after controlling for endogeneity of drug use and selection into employment. However, he only looks at recent cannabis use, while past use of cannabis has no significant effect.[7] Perhaps the clearest finding emerging from the literature on cannabis use and labour market outcomes is the co-existence of contradictory findings emerging from analyses based on the same datatsets (see, for example, the discussion by Gill et al. [1992], where they mention the need of "a more thorough understanding of the mechanism by which the positive relationship between drug use and earnings arises"). Our findings also shed some light on the apparent incompatibility of these different results. We point out that previous work has either neglected the role played by cognitive ability (mainly because of data limitations), or it has not properly taken it into account, simply including proxies such as test scores in the wage equation, rather than specifying cognitive ability as a latent vector with test scores as multiple indicators of an underlying trait (see Heckman et al. [2006]). Indeed, even linear fixed-effects models are unable to deal with unobserved heterogeneity appropriately, as they make strong assumptions on both the nature of the unobservables (a time-invariant factor is assumed to have the same effect in all periods) and the way in which they enter the model (i.e. additively).

3 Data: the British Cohort Study (BCS70)

The BCS70 is a longitudinal study which has followed the lives of a cohort of all children born in Great Britain in the week April 5-11, 1970. Originally designed to study perinatal

variables. The only comparable data source is the NLSY, for which information on drug use is available for the period 1979-1998, when the respondents were in the age range of 14-22 to 33-41. An advantage of the NLSY is that questions on drug use are asked at more points in time (1984, 1988, 1992 and 1998, and retrospectively from 1979). However, our data are much richer in terms of the amount of information they provide in relationship to drug-related knowledge, information, and perceptions, and in terms of the background variables.

[6] There are also studies on the employment effect of drug use, which are not mentioned here for expositional convenience.

[7] We are able to replicate his findings on our data only when we restrict our attention to the sample of regular drug users at 34.

mortality, it has expanded over time: six follow-up studies were conducted at the ages of 5, 10, 16, 26 (Ekinsmyth et al. [1992]), 30 (Bynner et al. [2000]) and 34 (Dodgeon et al. [2006]). In the birth survey, information was mainly provided by the midwife, and supplemented by clinical records. At age 5, 10 and 16, the cohort members undertook ability tests, and additional information was provided by the parents, who were interviewed by Health Visitors, and by the school health service; in addition, head and class teachers provided supplementary information when the cohort members were aged 10 and 16. Unfortunately, a low response rate occurred in the latter sweep because of a teacher strike during the fieldwork; however, analysis undertaken by Shepherd [1993] has shown that all school children were affected in the same way, and the sample remains representative of the population. The last two follow-ups were conducted by means of a mail survey (at age 26) and of direct interview (age 30). The response bias has been analysed in various occasions (Butler et al. [1997], Davie et al. [1972], Shepherd [1993]); it seems that the achieved samples do not differ from the target sample across a wide range of variables (especially social class, parental education, and gender), but the most disadvantaged groups are slightly underrepresented.

3.1 Drug Questions

The BCS70 questionnaire contains questions on cannabis in three different sweeps, widely spaced in time, allowing cannabis histories to be analysed for a very long period, from ages 16 to 34. A great advantage of our data is that we do not have to rely on recall questions on drug histories, thus avoiding the measurement error bias likely to be present in studies which make use of retrospective information.[8]

At age 16, the questions on drug use were included in the "Student self-completion Health Questionnaire", which was filled by the cohort member at the time of the medical examination, and returned to the doctor or nurse before leaving the place. A randomised list device was used to guarantee anonymity, after the following question:

> The next question [...] asks whether or not you have tried a number of substances some of which would under some circumstances be against the law. These are mixed in with a number of sporting activities and we have scrambled these by putting them into two lists - list A and list B. Please look at the box on this page to see whether you are to use list A or B when answering questions [...] to [...]. Please memorise whether it is list A or B you are to use then erase the letter A or B with ink. Then proceed to use the list indicated for answering [...] to [...]. Remember that nobody except you and us will know which list you are using.

The list includes 16 items: 7 types of illicit drugs (glues and solvents, cannabis, uppers, downers, lsd, cocaine and heroin), the dummy drug semeron and 8 sports.

At age 30, CASI (Computer-Assisted Self-Interviewing) was used with the self-completion instrument. This includes the section "Use of illegal drugs", which contains questions on the use of 12 types of illegal substances (cannabis, ecstasy, amphetamines, lsd, popper, magic

[8] However, unfortunately the structure of the drug questions is such that cannabis consumption can only be modelled as a discrete choice: this might be problematic if addiction or gateway effects depend on the amount consumed.

mushrooms, cocaine, temazepan, ketamine, crack, heroin, methadone). The phrasing of the question seems to suggest greater indulgence towards certain behaviour:

> As you know, many people have experimented with drugs at sometime. Have you ever tried cannabis, also known as blow, draw, puff, grass, skunk, weed, black, hash or red seal?

The survey also includes 'catchall' questions to capture the drugs not listed.

At age 34, only the following two questions on the use of cannabis[9] were asked in the self-completion instrument (CASI):

> Q1: (Since you were last interviewed on [...],) Have you ever tried cannabis - also known as marijuana, dope, hash, spliff, skunk or weed?

If has tried cannabis:

> Q2: (Since you were last interviewed on [...],) How often have you smoked cannabis?

Before analysing the data, we should first consider the reliability of self-reported data on drug use. Typically, we expect some form of underreporting - Mensch and Kandel [1988], in their analysis on NLSY data, found that this is particularly likely to happen among light users, females and ethnic minorities. Hoyt and Chaloupka [1994] found that also survey conditions are important in determining individual reporting behaviour (for example, the presence of others during the interview might lead to underreporting). Even if this is a common problem in drug research, we are fortunate in that the last two sweeps of the BCS70 have used the CASI method, which has been proved to have some advantage over traditional interviewing techniques, especially in terms of guarantee of anonimity (McAllister and Makkai [1991]). Another important issue to consider is recanting behaviour: we are fortunate in that, compared with other sources of data on drugs for the UK, the BCS70 seems to suffer less of this problem (see Pudney [2007]).

The unconditional frequencies and prevalence rates at each sweep are reported in Table 1.[10] The low proportion of individuals declaring to have used cannabis does not come as a surprise, given the young age of the sample members, the year of the survey (1986), and the incriminatory tone of the question. The reported drug use in the other two sweeps confirms the usual finding in the literature that individuals tend to experiment with drug before their 30s (with half of the sample having tried cannabis at least once before 30), and then to quit spontaneously (see, for example, Gill and Michaels [1991a]), so that current prevalence in the mid-thirties is not as high as for alcohol and tobacco. It is worth noticing that the legislation concerning cannabis has been subject to significant changes between the last two periods: in January 2004, cannabis has been reclassified from class B to class C drug.

[9] Note that cannabis was depenalised in January 2004, and the fieldwork for this sweep took place between February 2004 and June 2005.

[10] Columns 2 and 3 use all observations available at each sweep with non-missing information on drug use, columns 4 and 5 use only those observations with non-missing information on drug use in all three sweeps. It is worth stressing that the prevalence rates are remarkably similar in both samples.

Looking at the completed sequences of cannabis use (Table 2) reveals that many of the users are experimenters, who have quit using the drug by age 30. Remarkably, of those who have declared having tried cannabis by 30, but not in the previous year, a significant proportion (5.85%) reveals occasional use at the age of 34, and a small one (0.65%) reveals regular use. This stresses the importance of having longitudinal data spanning a long time for uncovering fictitious quits which end up in relapse.

Splitting the sample by gender confirms the usual finding in the literature that women are more likely than men to abstain from the consumption of illicit substances (Mullahy and Sindelar [1991]) and that, if they try them, they are more likely to be experimenters. Interestingly, the women in the British sample do not seem to display the "multi-commodity habit formation" that Pacula [1997] found for US data. We find limited evidence that, whenever the females consume illicit substances, they develop a habit for multiple consumption: the proportion of males and females who have tried different numbers of substances is remarkably similar, which again remarks the similarity of experiences at this developmental stage.[11]

Table 1 Cannabis use[12]

Age 16	n	%	n	%
Never used	6,193	93.24	4,159	93.57
Used[13]	449	6.76	286	6.43
	6,642[14]	100.00	4,445	100.00
Age 30	n	%	n	%
Never used	5,470	49.36	2,247	50.55
Used, not in last 12 months	3,729	33.65	1,508	33.93
Used in last 12 months	1,882	16.98	690	15.52
	11,081[15]	100.00	4,445	100.00
Age 34	n	%	n	%
Non-user[16]	7,353	76.61	3.520	79.19
Occasional user[17]	1,735	18.08	724	16.29
Regular user[18]	510	5.31	201	4.52
	9,598[19]	100.00	4,445	100.00

Legend: n=Frequencies; %=Prevalence rates. See also footnote 10.

[11] For males, 65.95% report having used only one illicit drug, 20.00% two, 7.62% three, 4.76% four, 1.43% five and 0.24% six; the respective numbers for the females are 67.39% for one, 21.47% two, 7.61% three, 2.45% four and 1.09% five. The drugs considered here are cannabis, glues and solvents, upper and downers, lsd, cocaine and heroin. See Table A4 for cannabis use by gender.

[12] 55 individuals reporting the use of semeron (31 at 16 and 24 at 30) have been dropped from the sample.

[13] This groups the following four categories: *Yes, not in the past year [105]; Yes, once in the past year [149]; Yes, 2-9 times in the past year [137]; Yes, 10+ times in the past year [58].*

[14] Nonmissing observations. Missing observations at 16 are 9,864, 59.76% of the available sample (n=16,506).

[15] Nonmissing observations. Missing observations at 30 are 5,425, 32.87% of the available sample.

[16] Note the question refers to the time period since last interview.

[17] This groups the following four categories: *Two or three times a month [98]; Once a month [74]; Less often/only on special occasions [895]; I never use cannabis nowadays [668].*

[18] This groups the following three categories: *On most days [289]; Two or three times a week [156]; Once a week [65].*

[19] Nonmissing observations. Missing observations at 34 are 6,908, 41.85% of the available sample.

However, if we also consider the consumption of alcohol and cigarettes, females show a clear preference for double addiction (41.27% reports to have used 2 substances, while the proportion for males is 35.80%); however, this finding is mainly driven by the fact that, by age 16, 48.05% of the females have already tried a cigarette, while this proportion stops at 42.10% for males.

Table 2 Completed sequences of cannabis use

Type	Sequence	n	%
Abstainers	000	2,139	48.12
Middle Experimenters	010	1,071	24.09
Middle Recreational Users	021	291	6.55
Light Relapsers I	011	260	5.85
Quitters I	020	154	3.46
Middle Regular Users	022	131	2.95
Early Experimenters	110	119	2.68
Late Starters I	001	88	1.98
Early Recreational Users	121	56	1.26
Early Regular Users	122	39	0.88
Light Relapsers II	111	29	0.65
Heavy Relapsers I	012	23	0.52
Quitters II	120	19	0.43
Recanters	100	18	0.40
Heavy Relapsers II	112	6	0.13
Late Starters II	002	2	0.004
		4,445	100.00

Source: BCS70, ages 16, 30 and 34.
Legend for the sequences - 1st col.: 0=Never used by 16, 1=Used by 16;
2nd col.: 0=Never used by 30, 1=Used more than 12 months ago,
2=Used in last 12 months at 30; 3rd col.: 0=Never used between 30 and 34,
1=Occasional user at 34, 2=Regular user at 34.
See Table A2 for fragmented sequences.

3.2 Cognitive Ability Tests

At age 5 five tests were administered, but only four are used, because the Schonell's reading tests was not administered to 7,422 children. These are the Copying Design Test, the Human Figure Drawing Test, the English Picture Vocabulary Test and the Profile Test. The Copying Designs Test assesses the cohort member's visual-motor coordination (Davie et al. [1972], Rutter et al. [1970]). The children were asked to make two copies of each of the 8 designs reproduced in the Test Booklet. The Human Figure Drawing Test used in the BCS was a modified version of the Draw-a-Man Test originally created by Goodenough [1926] and further developed by Harris [1963], and assesses the child's 'conceptual maturity'. The child was asked to 'make a picture of a man or a lady', and another picture of a person the opposite sex to the first. The scoring scheme was based on 30 items of a developmental scale suggested by Koppitz [1968], using the Harris [1963] scoring system. The English Picture Vocabulary Test is an adaptation of the American Peabody Picture Vocabulary Test (PPVT) by Brimer and Dunn [1962]. It consists of 56 blocks of 4 different pictures, with a word associated with each block of 4. The child is asked to indicate the picture corresponding to the given word, and the test continues with words of increasing difficulty until the child makes five mistakes in 8 consecutive items. The Profile Test was used with the permission of his author Kalverboer [1972], and assesses the child's spatial-constructive development. The child was asked to complete the profile drawn on the Test Booklet. The score was then based on the number of correct facial items drawn by the child.

At age 10, four tests were administered: the British Ability Scales, the Shortened Edinburgh Reading Test, the Friendly Math Test and the Pictorial Language Comprehension Test.[20] The British Ability Scales is a test measuring a mental construct similar to IQ (Elliot et al. [1978]). It comprises two verbal scales: word definitions (37 items) and word similarities (21 items) and two non-verbal scales: recall of digits (34 items) and matrices (28 items). The Edinburgh Reading Test is a test of word recognition; in the BCS70 a shortened version was used, upon consultation with its authors (Godfrey Thomson Unit [1978]). This contains 75 items to examine vocabulary, syntax, sequencing, comprehension and retention, and is particularly designed to capture poor readers. The Friendly Math Test was specifically designed for the British Cohort Study, given the unavailability, at the time of the survey, of a suitable test for 10-year olds. It contains 72 items testing understanding of the rules of arithmetic, fractions, algebra, geometry and statistics. Finally, the Pictorial Language Comprehension Test is also a new test specifically developed for the BCS70 on the basis of the American Peabody Picture Vocabulary Test and its English counterpart, the English Picture Vocabulary Test. It contains 71 items covering vocabulary, sequence and sentence comprehension. The BCS70 questionnaire at age 16 (Youthscan) includes five tests of cognitive ability: reading, matrices, arithmetic, vocabulary and spelling. Details on all the tests available in the BCS70 are included in Table A1.

In our analysis, we do not use the tests administered when the child was aged 16, to avoid any possible source of simultaneity with contemporaneous use of cannabis. Interestingly, the relationship between late indicators of cognitive ability (i.e. the tests administered at age 16) and cannabis use seems to be less in line with the positive association observed using earlier indicators (i.e. the tests administered at age 5 and 10), and more consistent with a negative association as found in other studies. This is true both when pooling all the observations (see Table 3) and when looking at each gender separately (see Table A3). This observation is actually suggestive of a causal effect of cognitive ability on cannabis use: it seems that what matters is how early (or how fast) you develop, not what score you are able to get - also, bear in mind that the scores at early age are more likely to reflect innate factors than acquired skills.

[20]In the analysis, we have disaggregated the BAS into the four sub-components, as they capture different cognitive domains.

Table 3 Cognitive ability by cannabis use

Test	Age 16		Age 30			Age 34		
	Never used	Used	Never used	Used not last 12m	Used last 12m	Non user	Occas. user	Regular user
Age 5								
CDT	0.614	0.669	0.575	0.637	0.630	0.605	0.639	0.587
HFDT	0.471	0.504	0.460	0.477	0.480	0.469	0.479	0.462
EPVT	0.651	0.716	0.617	0.666	0.671	0.641	0.677	0.645
PT	0.435	0.470	0.420	0.446	0.458	0.430	0.454	0.447
Age 10								
PLCT	0.581	0.600	0.561	0.590	0.593	0.576	0.597	0.576
FMT	0.638	0.691	0.599	0.639	0.651	0.624	0.654	0.621
SERT	0.644	0.712	0.593	0.644	0.657	0.624	0.658	0.611
BAST-M	0.581	0.631	0.543	0.582	0.591	0.568	0.597	0.550
BAST-S	0.591	0.637	0.563	0.598	0.611	0.580	0.612	0.592
BAST-WD	0.291	0.344	0.258	0.300	0.316	0.279	0.314	0.290
BAST-RD	0.666	0.694	0.651	0.673	0.678	0.661	0.678	0.661
Age 16								
VOC	0.588	0.607	0.557	0.587	0.599	0.574	0.604	0.572
SPEL	0.820	0.805	0.813	0.821	0.814	0.820	0.822	0.795
MATH	0.638	0.652	0.605	0.639	0.633	0.628	0.637	0.580

"Used at 16" groups the following four categories: *Yes, not in the past year; Yes, once in the past year; Yes, 2-9 times in the past year; Yes, 10+ times in the past year.*

"Non user at 34": note the question refers to the time period since last interview.

"Occasional user at 34" groups the following four categories: *Two or three times a month; Once a month; Less often/only on special occasions; I never use cannabis nowadays.*

"Regular user at 34" groups the following three categories: *On most days; Two or three times a week; Once a week.*

4 Preliminary Evidence

We first report some OLS evidence on the basic associations present in the data. Table 4 shows OLS estimates of the effect of cognitive ability at age 5[21] on three indicators of cannabis use (Panel A)[22], and of the effect of cannabis use on wages (Panel B). First, I find a large and significant association between cannabis use and the test scores, which is remarkably stable across specifications and basically unaffected by the inclusion of a large set of controls for background characteristics. Second, I document a positive association between wages at 34 and lifetime prevalence and experimentation with cannabis use by age 30, which is still present upon the inclusion of background characteristics, but greatly shrinks in magnitude and it is driven to insignificance after controlling for cognitive ability.

[21] We use the four tests at age 5 described in the previous section.

[22] The three indicators are: early initiation (by 16), lifetime prevalence by 30 and experimentation by 30.

Table 4 Cannabis use, cognitive ability and wages

Panel A. Cannabis use and cognitive ability

Dep. Var.	*Test at 5*	*Model 1*	*Model 2*
Cannabis use by 16	CDT	-0.300	-0.320
		(0.187)	(0.198)
	HFDT	1.141	1.208
		(0.394)	(0.407)
	EPVT	1.090	1.065
		(0.266)	(0.293)
	PT	0.156	0.172
		(0.165)	(0.172)
Cannabis use by 30	CDT	0.137	0.106
(life prevalence)		(0.094)	(0.099)
	HFDT	0.643	0.674
		(0.206)	(0.211)
	EPVT	0.692	0.670
		(0.124)	(0.133)
	PT	0.091	0.096
		(0.084)	(0.086)
Cannabis use by 30	CDT	0.203	0.203
(experimentation)		(0.094)	(0.099)
	HFDT	0.335	0.347
		(0.204)	(0.208)
	EPVT	0.412	0.414
		(0.215)	(0.133)
	PT	-0.018	-0.004
		(0.083)	(0.085)

Panel B. Cannabis use and wages

Dep. Var.	*Cannabis*	*Model 1*	*Model 2*	*Model 3*
Wage at 34	use by 16	0.065	0.078	0.040
		(0.049)	(0.055)	(0.057)
	life prevalence	0.051	0.035	0.012
		(0.017)	(0.018)	(0.019)
	experimentation	0.056	0.047	0.030
		(0.017)	(0.019)	(0.019)

The sample includes only males. Model 1 in Panel A only includes the test scores, Model 2 also includes indicators for parental education, parity, mother's age at birth and paternal social class. Model 1 in Panel B only includes cannabis use, Model 2 also includes the same family background indicators as in Panel A, Model 3 also includes the four test scores administered at age 5.

I now turn to the description of the econometric model I will use in the remainder of the paper, in order to take into account measurement error in test scores and fully exploit the longitudinal structure of the data.

5 Econometric Analysis of Ability and Drug Use

We observe drug use in three periods: when the individual is aged 16, at 30 and at 34. The rich longitudinal structure of our data allows us to overcome the usual simultaneity problem that is faced by researchers when looking at outcomes jointly determined with drug use: cognitive ability in our data is measured at a very early stage of the life, thus anticipating any possible experience of drug use. Nonetheless, we face the usual heterogeneity problem: unobserved characteristics influencing the individual's choice of taking drugs could well influence the process of cognitive development. Our hypothesis is that a high level of intellectual curiosity will be conducive both to early cognitive development and to experimentation with different sorts of forbidden behaviour, including drug use.

To address all these issues, we estimate a joint model of drug use across the lifecycle together with cognitive ability.

5.1 Ability measurement

Ideally, what we would like to have is a small number of measures of cognitive achievement which reflect the capabilities of the adolescent mind in different domains. Conventionally, such measures have been created using principal component analysis or similar techniques aiming at reducing data dimensionality. However, we have shown elsewhere (Conti and Pudney [2007]) that the PCA approach suffers of serious shortcomings if used to generate ability measures in this way. Instead, for simplicity here we assume that there is only one cognitive ability (see Conti and Pudney [2007] for a generalisation).

We first need to specify a *measurement model* for the test scores. We assume that the test scores are independently normally distributed conditional on true abilities. For each test j, there is a linear testing technology, relating the test scores to the underlying cognitive ability q of subject i:

$$\tau_{ij} = \phi_j + \alpha_j q_i + v_{ij} \tag{1}$$

where α_j is the loading on test j, and v_{ij} is a purely stochastic measurement error component of test performance, assumed to be normally distributed with zero mean and $var(v_{ij}) = \sigma_j^2$. We model cognitive ability using the four tests administered to the child when was aged 5 (so $j = 1, 2, 3, 4$), to allow the longest possible time between indicators of cognitive capabilities and subsequent patterns of behaviour, and to minimise the role played by compensating/reinforcing factors affecting the innate abilities of the child.[23] The specification of the cognitive factor is assumed to be the following linear regression:

$$q_i = \mathbf{Z}_i\boldsymbol{\theta} + u_i \tag{2}$$

where $u_i \sim N(0, \omega^2)$, $u_i \perp v_{ij}$ and $\mathbf{Z}$ contains basic background factors observed at the time of the birth (such as parental education, age and social class) which capture heterogeneity in cognitive endowments and are also allowed to affect drug use directly. It is well known that in factor models at least one factor loading or the factor variance must be fixed to a nonzero constant, otherwise the scale of the factor is not identified. In the following we adopt two

[23]In ongoing work, we experiment using the tests administered at age 10. In addition to capture different abilities as compared to the tests administered at 5, the comparison of the results would allow us to shed some light on the role played by the timing and the speed of development.

different normalisations: in the sequence analysis we set the variance of the factor to one and estimate all the loadings; in the analysis of sequential behaviour we set the loading on the copy design test in the measurement model to one.

5.2 Drug consumption

We adopt two alternative approaches for the specification of the part of the model which links drug use to cognitive ability.

5.2.1 Sequence analysis

As a first approximation, we group the 15 observed sequences $y = \{y_1, y_2, ...\}$ into 4 classes,[24] on the assumption that $\Pr(\text{type } k|y \in S_k) = 1$, where S_k is an *a priori* class of realisations of $\{y_1, y_2, ...\}$, and y_t represents drug use in period t. The basic relationship between cognitive ability and trajectories of drug use is shown in Table A6: what emerges at first glance is that sequences of experimentation and moderate drug use behaviour are associated with higher scores on tests of cognitive ability, especially at younger ages.[25] Splitting the sample by gender (see Table A5) reveals that there are no basic differences in the associations between sequences of drug use and cognitive ability by males and females, which allows us to pool the sample by gender.

The probability of following trajectory j is then modeled as a multinomial logit:

$$\Pr(j|\mathbf{X}, q) = \frac{exp(\mathbf{X}\boldsymbol{\beta_j} + q\gamma_j)}{\sum_k exp(\mathbf{X}\boldsymbol{\beta}_k + q\gamma_k)} \tag{3}$$

where $\mathbf{X}$ is a vector containing both background factors and events happening up to age 16, and β_0 is normalised at 0.

The elasticity of the expected number of drug users of type j with respect to cognitive ability q is computed in the following way:

$$\frac{\partial \ln \Pr(j|\mathbf{X}, q)}{\partial \ln q} = \left[\beta_j - \sum_{k=0}^{3} \Pr(k|\mathbf{X}, q)\beta_k\right] q \tag{4}$$

The standard errors have been computed using the delta method.

[24]See Tables 2 and A5. The sequences are 15 because the “recanters” are dropped from the analysis. Then, the other categories are grouped in the following way: the “abstainers” constitutes the baseline category; the “experimenters” group the categories “early experimenters”, “middle experimenters”, “quitters I” and “quitters II”; the “recreational users” group the categories “late starters I”, “light relapsers I”, “middle recreational users”, “light relapsers II” and “early recreational users”; finally, the “regular users” group the categories “late starters II”, “heavy relapsers I”, “middle regular users”, “heavy relapsers II” and “early regular users”. Of course it would be ideal to analyse more disaggregated categories, unfortunately sample size issues hinder us from exploiting this possibility.

[25]Surprisingly, these patterns are completely the opposite as shown for US data by, for example, Pacula et al. [2003], where the abstainers reported consistently higher test scores, and experimenters lower.

5.2.2 Simultaneous estimation

As a second approach, we specify a *disease model* which links the determination of drug use to cognitive ability. We define a latent variable d_i^* which represents an individual's unobserved propensity to consume cannabis over some specific period. The observable counterparts of these latent variables are binary indicators of drug use, which are generated according to the following usual probit mechanism:

$$d_i^* = \mathbf{X}_i\boldsymbol{\beta} + q_i\gamma + \varepsilon_i \tag{5}$$

$$d_i = 1(d_i^* > 0) \tag{6}$$

where 1(.) is the indicator function, which equals one if the argument is true and zero otherwise, and ε_i is a random error which is normally distributed, conditional on the covariates. A separate relationship of the form (3.5) is specified for each of the periods covered by successive waves of observation.

The likelihood function then is:

$$L = \prod_i \int \Phi\left(\mathbf{X}_i\boldsymbol{\beta} + q_i\gamma\right)^{d_i}\left[1 - \Phi\left(\mathbf{X}_i\boldsymbol{\beta} + q_i\gamma\right)\right]^{(1-d_i)} \times \left[\prod_{j=1}^{4} \frac{1}{\sigma_j}\phi\left(\frac{\tau_{ij} - \phi_j - \alpha_j q_i}{\sigma_j}\right)\right] \tag{7}$$

$$\times \frac{1}{\omega}\phi\left(\frac{q_i - \mathbf{Z}_i\boldsymbol{\theta}}{\omega}\right) dq_i$$

where $\phi(.)$ is the pdf of the $N(0,1)$ distribution.

5.3 Estimates of the ability/drugs relationship

The first set of estimates refers to the multinomial logit analysis of the effect of cognitive ability on the probability of following a certain drug trajectory. Statistics for the sample used are reported in Table A8. They show, not surprisingly, that males are more likely to follow patterns of heavier drug use as compared to females.[26] However, we notice that individuals with histories of experimentation or experience of moderate drug use come from wealthier backgrounds, with parents more educated and from a higher social class. The abstainers are over-represented among the lowest social class, while those individuals who end up using cannabis regularly by the age of 34 are more likely to have been born to lone mothers, not as first child, and to have experienced a dramatic situation in the critical period of adolescence, such as parental divorce, illness or disability, or a personal accident.[27]

Estimated coefficients for the sequence analysis are reported in Table A9 ("abstainer" is the baseline category). However, since these coefficients are not directly interpretable, we concentrate our attention on the elasticities for specific covariates, which we compute using equation (4) above, for two different types of individuals. The choice of these two types is mainly driven by the interest in comparing an "average" individual with an individual coming from a situation of disadvantage, both in terms of family background factors and

[26]Note that, differently from the rest of the analysis, the sample here includes both males and females.

[27]Note that we use information only on experiences occurred between the ages of 10 and 16 not only because of this being a critical period, but also to avoid any possible source of simultaneity with the test scores, by allowing (at least) a 5-period lag between them and the experiences recorded here.

because of adverse experiences during adolescence. Firstly, we consider a female born as the first child from a 25-year old mother and a father belonging to the 3rd manual social class, with a younger sibling and both parents with only the compulsory level of schooling, no adverse experiences during childhood and living in a region with a moderately low level of drug seizures (0.4).[28] Our second simulation is carried out for a male born to a 25-year old mother who already had other two children and a low social class father (as before, both parents with no education beyond the compulsory level), who experienced both parental divorce and an accident during adolescence, while living in a region with a drugs market of medium dimensions (the value we have chosen for the seizures is 0.5 in this case).

Table 5 Predicted probabilities and elasticities

	Abstainer	**Experimenter**	**Recreational user**	**Regular user**
Predicted probabilities				
Type 1	0.295	0.239	0.384	0.082
Type 2	0.040	0.117	0.266	0.576
Elasticities - Type 1				
Ability	−0.281	0.148	0.124	−0.002
	(0.065)	(0.057)	(0.054)	(0.128)
Seizures	−0.583	0.208	0.257	0.283
	(0.175)	(0.151)	(0.148)	(0.443)
Mother's age	1.637	−2.018	0.507	−2.355
	(0.175)	(0.151)	(0.148)	(0.443)
Elasticities - Type 2				
Ability	−0.206	0.071	0.056	−0.026
	(0.071)	(0.060)	(0.061)	(0.040)
Seizures	−0.526	−0.016	0.016	0.032
	(0.222)	(0.199)	(0.203)	(0.131)
Mother's age	1.954	−0.403	1.226	−0.619
	(0.222)	(0.199)	(0.203)	(0.131)

Standard errors in parentheses.

First of all, we notice (see Table 5) that our model predicts, as expected, different trajectories of use for the two different types, with the second one clearly more likely to be prone to addiction. Turning to the effect of the observed covariates, an interesting finding clearly emerges: a 1% increase in the quantities of controlled drugs seized by the police only reduces the probability of being an abstainer, but it has no significant effect on the probability of falling into one of the categories of drug users. This has a clear policy relevance: drugs availability only seems to affect the likelihood of trying drugs, not the subsequent pattern of use, which is, apparently, mostly determined by personal characteristics. In particular, having a teenage birth seems to affect adversely the life chances of both types, but on the consequences of early childbearing there is already a consensus in the literature. Remarkably, the "thickness" of the drug market seems to act independently of the type, while the effect of cognitive ability clearly differs depending on the particular circumstances the individual is born in.[29] Indeed, if more able individuals are less likely to be abstainers, their cognitive potential predisposes them towards relatively harmless patterns of behaviour if their initial

[28]Indicator of drug prevalence as measured by the number of controlled drugs seized by the police per head of population.

[29]This is a facet of the ever-present debate of the relative role played by genes *vis-a-vis* environment and of the interaction between the two.

conditions do not put them in a situation of extreme disadvantage,[30] while it has no significant effect if they already come from a situation of extreme disadvantage. As the individuals with a more dramatic pattern of drug-related behaviour are more easily detectable on the basis of personal characteristics, it is our belief that drug-related policy should be targeted at these particular group of individuals, rather than being on a universal basis.

The importance of health education versus economic/legal constraints is reflected in the answers given to other questions about the (lack of) use of substances. Remarkably, the majority of the respondents declare not having tried drugs, not because of lack of opportunity, but as a matter of choice. This might have important policy consequences: if the initiation process is mostly driven by non-economic factors, the policy debate should shift from focusing on the legal aspects to increasing awareness of the health consequences of substance use in the longer term.

What emerges from the sequence analysis is clear: cognitive ability is a strong and significant predictor of certain patterns of drug use, and this effect varies by background conditions. In order to analyse this relationship in more depth, we now turn to the analysis of sequential behaviour, where we look at the effect of cognitive ability on drug use separately for two different ages (16 and 30),[31] and we restrict our attention to a male-only sample of employed individuals.[32] We define drug use at 30 in two different ways to try to overcome the problem of "partial observability" inherent in the phrasing of the drug question: we first define a binary indicator for "ever used drug by 30", which captures both those who have quit and those who are still using, and then we define another binary indicator for "experimenter", which captures those who report at 30 having used cannabis more than 12 months ago. Summary statistics by type of user are reported in Table A10. We first turn our attention to the first two columns and see that individuals who reported having use cannabis by the age of 16[33] exhibit a combination of two different sets of characteristics: on the one side, they come from relatively affluent backgrounds; on the other, they are more likely to have experienced difficult situations such as parental divorce and/or father with an illness or a disability. The estimated coefficients (Table A11, column 1) show clearly that there are two basic sets of factors affecting early initiation: one is related to family circumstances occurring during adolescence (having older siblings and a father being ill or disabled), the other is linked to the "thickness" of the drug market.[34]

Now turn to the determinants of drug use by 30, and consider columns 3-4-6 of Table

[30]The marginal effects of ability on the probability of being both an experimenter and a recreational user are both positive and significant for Type 1, though not big in magnitude.

[31]We do not use the information we have on drug use at 34 in order to avoid bias due to simultaneity between drug use and wages. Also, there might be confounding effects due to changes in individual behaviour (either in terms of reporting, or in terms of actual consumption of drugs) as an effect of the depenalisation of cannabis occurred in 2004.

[32]We leave out self-employed for now as drug use is clearly related to occupational choice (see Fairlie [2002]) and do not deal with the issue of selection into employment now to simplify matters. However, we believe it is less of an issue in case of a male sample, which is the main reason why we do not consider females in the remainder of this study.

[33]Note that the small sample size here is a combination of different things: it is basically a mix of the low-response obtained in the BCS70 for the 3rd sweep, of the under-reporting due to the "incriminatory" tone of the survey question, and of a true low incidence of drug use at young age in Britain in the 1980s.

[34]The latter finding has recently been highlighted in the literature: see, for example, Van Ours and Williams [2007] for a positive effect of low cannabis prices on early onset. Also see Kandel [1980] for the lack of any effect of social class on drug use.

A10.[35] As we have seen already, individuals who have tried cannabis by the age of 30 are both more likely to come from affluent backgrounds and to have experienced dramatic circumstances in the critical period of adolescence (in addition to having being exposed to greater drug availability). Noticeably, these individuals have also stayed in education beyond the compulsory level,[36] and they earn higher wages. When we turn to the estimated coefficients (Table A11, cols. 2 and 3), we see that factors affecting lifetime prevalence and/or experimentation differ from those determining early initiation. Firstly, there are events occurring during adolescence which have a long-term effect: parental divorce, the loss of a job by the father and the experience of an accident all seem to have a delayed impact, affecting lifetime prevalence. Secondly, the coefficient on seizures is greatly reduced in magnitude, and significant at 10% level only in the model for drug prevalence; however, as early initiation is correlated with habit formation (see Pudney [2004] and Van Ours and Williams [2007]), we cannot rule out the possibility that this is the case. Noticeably, there are almost no observable factors predicting experimentation, if we leave out cognitive ability.

6 Cannabis, Ability and Wages

In the previous section we have analysed the relationship between cognitive ability and drug use, looking at sequences of behaviour and joint modeling. Three basic findings have emerged: cognitive ability is significantly related to cannabis use by 30; the "thickness" of the drugs market plays a greater role in the early initiation process; family-related events occurring during adolescence are more important than background factors in determining patterns of drug use, and they operate with different timing. As it seems that drug use is associated with factors which can carry both a wage premium and a wage penalty,[37] it is natural now to turn to the analysis of the effect of cannabis use on wages. In particular, our aim is to try to understand whether there is an independent effect of the former on the latter, or any effect is driven by both observable and unobservable factors correlated with both.

6.1 The model with earnings

We now introduce earnings, and we firstly estimate a simple endogenous dummy variable model, in order to compare our findings with those in the literature. So, we estimate jointly the probability of drug use and the process of wage determination:

$$d_i^* = \mathbf{X}_i\boldsymbol{\beta} + \varepsilon_i \tag{8}$$

$$w_i = \mathbf{W}_i\boldsymbol{\delta} + q_i\lambda + d_i\psi + \eta_i \tag{9}$$

$$\begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma_\omega \\ \rho\sigma_\omega & \sigma_\omega^2 \end{bmatrix}\right) \tag{10}$$

[35]Column 5 is not of independent interest, as it groups individuals who have both never tried cannabis by 30, and who have are still using it.

[36]It is interesting to notice also that, for what concerns marital status, there is a clear difference across types, with a low marriage rate among individuals who are still using cannabis at 30.

[37]As highlighted in the literature review, every finding is possible.

To ensure identification, we exclude from the wage equation factors believed to be associated only with the use of illicit drugs, such as religious affiliation and the quantity seized in the Police Force Area in which the individual lived at the time when he was 16. In addition to this, we exploit the longitudinal structure of our data and we enter in the wage equation life episodes happened in periods subsequent to those the information on cannabis use refers to.

Finally, we introduce the cognitive factor, and we build up the following likelihood function:

$$L = \prod_i \int \frac{1}{\sigma_\omega}\phi\left(\frac{w_i - \mathbf{W}_i\boldsymbol{\delta} - q_i\lambda - d_i\psi}{\sigma_\omega}\right)\Phi\left(\mathbf{X}_i\boldsymbol{\beta} + q_i\gamma\right)^{d_i}\left[1-\Phi\left(\mathbf{X}_i\boldsymbol{\beta}+q_i\gamma\right)\right]^{(1-d_i)} \quad (11)$$
$$\times\left[\prod_{j=1}^{4}\frac{1}{\sigma_j}\phi\left(\frac{\tau_{ij}-\phi_j-\alpha_j q_i}{\sigma_j}\right)\right]\times\frac{1}{\omega}\phi\left(\frac{q_i-\mathbf{Z}_i\boldsymbol{\theta}}{\omega}\right)dq_i$$

6.2 Estimates of the ability/drugs/earnings relationship

We have estimated separate models for the processes of early initiation into drug use, lifetime prevalence by 30 and experimentation by 30, jointly with the process of cognitive development at the age of 5 and wage determination at the age of 34. The results are shown in Tables A12-A14. For each table, we present two sets of results: in the first column we abstract from the process of cognitive development, and only look at the relationship between drug use and wages, while controlling for other unobserved factors which could be linked to both, as has been done in the literature up to this stage (see, for example, Kaestner [1991]). We first notice that the specification of the drug-use part of the model replicates the basic findings already highlighted in the previous section, and that all the variables in the wage equations have the expected sign. We replicate the main effects of interest in Table 6 for expositional convenience.

Table 6 Cognitive Ability, Drug Use and Wages

Dependent Variable: log hourly net wage at 34		
Ever used cannabis by 16	0.216 (0.170)	0.053 (0.057)
Ability	-	0.374*** (0.110)
Ever used cannabis by 30	0.603*** (0.049)	0.0295 (0.020)
Ability	-	0.348*** (0.086)
Experimented before 30	0.587*** (0.053)	0.0325 (0.020)
Ability	-	0.349*** (0.085)

Standard errors in parentheses.

Column 2: Model without cognitive ability.

Column 3: Model with cognitive ability.

The finding is quite simple and striking: while there is no effect of early initiation into cannabis on wages, the positive and significant effect which emerges in the two models for lifetime prevalence and experimentation by the age of 30 completely disappears once we control for cognitive ability. Indeed, the positive and significant effect of the ability factor (as measured at age 5) on wages is remarkably stable across the different specifications for the use of drugs. This is even more important if we notice that all the other coefficients are remarkably stable across the two columns.

This finding casts serious doubts on all the previous studies which have looked at the relationship between drug use and wages, and found significant effects in one direction or another. After all, it is reasonable to think (as noticed in medical studies) that the "impairment" effect of drug use on wages would only arise from problematic and heavy patterns of behaviour, mainly experienced by individuals who are likely to be under-represented in a general population sample. Also, in case of drugs used since young, this effect would mainly work through human capital formation, as noted in recent papers by Bray [2005] and Ringel et al. [2006]. In addition to this, our findings raise serious concerns about the use of identification strategies which are based on instruments related to family characteristics or parental behaviour, as they are likely to be correlated with unobservables also affecting cognitive ability, posing some threat to instrument validity.[38]

7 Conclusions

This paper has revisited the effect of cannabis use on wages by introducing cognitive ability as a previously unexplored channel accounting for the unobserved factors linking the two recognised in the literature so far. It has used a rich, unexplored dataset containing information covering test scores, drug use and wages at three very widely spaced time periods, overcoming the simultaneity problem present in most of the previous studies. It has provided the first set of estimates of this relationship for the United Kingdom, where drug policy has been recently again hotly debated. We have provided convincing evidence of the existence of a positive and significant relationship between a cognitive ability factor and lifetime cannabis use, which accounts for all the relationship between drug use and wages, and has been omitted in previous studies.

[38]DeSimone [2002] points out serious shortcomings in the use of the conventional instruments adopted in the literature on drugs and the labour market, and stresses the need of carefully testing for weak instruments and instruments validity. Unfortunately, he also finds that prices and the drug decriminalisation laws are much weaker instruments than indicators of parental background, which might suffer of the problem outlined above.

References

[1] Becker, G.S. and K.M. Murphy (1988), "A Theory of Rational Addiction", *Journal of Political Economy*, 96(4), 675-700.

[2] Bray, J.W. (2005), "Alcohol Use, Human Capital, and Wages", *Journal of Labor Economics*, 23(2), 279-312.

[3] Brimer, M.A. and L.M. Dunn (1962), *English Picture Vocabulary Test*, Bristol: Education Evaluation Enterprises.

[4] Butler, N., S. Despotidou and P. Shepherd (1997), *1970 British Cohort Study (BCS70) ten-year follow-up: A guide to the BCS70 10-year data available at the Economic and Social Research Unit data archive*, London: City University, Social Statistics Research Unit.

[5] Bynner, J., N. Butler, E. Ferri, P. Shepherd and K. Smith (2000), "The Design and Conduct of the 1999-2000 Surveys of the National Child Development Study and the 1970 British Cohort Study", *Centre for Longitudinal Studies Working Paper 1*, London: Institute of Education.

[6] Chen, S. (2006), "Does College Teach Young Men to Smoke Pot?", State University of New York at Albany, *mimeo*.

[7] Conti, G. and S.E. Pudney (2007), "The Dynamics of Cognitive Development", paper presented at the *2007 EALE Conference*.

[8] Conti, G., J.J. Heckman and S. Urzua (2010), "The Education-Health Gradient", *American Economic Review Papers and Proceedings*, forthcoming.

[9] Culyer, A. (1973), "Should social policy concern itself with drug abuse?", *Public Finance Quarterly*, 1, 449-456.

[10] Davie, R., N.R. Butler and Goldstein (1972), *From Birth to Seven. A Report of the National Child Development Study*, London: Longman.

[11] DeSimone, J. (2002), "Illegal Drug Use and Employment", *Journal of Labor Economics*, 20(4), 952-977.

[12] Dodgeon, B., J. Elliott, M. Hancock, J. Johnson, S. Parsons and P. Shepherd, *British Cohort Study (1970) 2004-2005 Follow-up Deposited with the UK Data Archive. A Guide to the Dataset*, Centre for Longitudinal Studies, Bedford Group for Lifecourse & Statistical Studies, Institute of Education, University of London.

[13] Fairlie, R.W. (2002), "Drug Dealing and Legitimate Self-Employment", *Journal of Labor Economics*, 20(3), 538-567.

[14] Gill, A.M. and R.J. Michaels (1991), "The determinants of illegal drug use", *Contemporary Economic Policy*, 9(3), 93-105.

[15] Gill, A.M. and R.J. Michaels (1991), "Does Drug Use Lower Wages?", *Industrial and Labor Relations Review*, 45(3), 419-434.

[16] Gill, A.M., R.J. Michael, C.A. Register and D.R. Williams (1992), "Employment and Earnings Effects of Drug Use. Discussion by the Authors", *Industrial and Labor Relations Review*, 45(3), 449-451.

[17] Goodenough, F. (1926), *Measurement of Intelligence by Drawings*, New York: Harcourt, Brace and World.

[18] Harris, D.B. (1963), *Children's Drawings as Measures of Intellectual Maturity*, New York: Harcourt, Brace and World Inc.

[19] Heckman, J.J., J. Stixrud and S. Urzua (2006), "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior", *Journal of Labor Economics*, 24(3), 411-482.

[20] Heyman, G.M. (2009), *Addiction. A Disorder of Choice*, Harvard University Press.

[21] Hoyt, G.M. and F.J. Chaloupka (1994), "Effect of survey conditions on self-reported substance use", *Contemporary Economic Policy*, 12, 109-121.

[22] Kaestner (1991), "The Effect of Illicit Drug Use on the Wages of Young Adults", *Journal of Labor Economics*, 9(4), 381-412.

[23] Kaestner (1994), "New Estimates of the Effect of Marijuana and Cocaine Use in Wages", *Industrial and Labor Relations Review*, 47(3), 454-470.

[24] Kalverboer, A.F. (1972), *A profile test for the spatial-constructive development*, Lisse: Switz and Zeitlinger.

[25] Kandel, D.B. (1980), "Drug and Drinking Behavior among Youth", *Annual Review of Sociology*, 6, 235-285.

[26] Kolb, L. (1925), "The relation of intelligence to the etiology of drug addiction", *American Journal of Psychiatry*, 82, 163-167.

[27] Koppitz, E.M. (1968), *Psychological evaluation of Children's Human Figure Drawings*, New York: Grune and Stratton.

[28] MacDonald, Z. and S. Pudney (2000), "Illicit drug use, unemployment, and occupational attainment", *Journal of Health Economics*, 19, 1089-1115.

[29] MacDonald, Z. and S. Pudney (2001), "Illicit drug use and labour market achievement: evidence from the UK", *Applied Economics*, 33, 1655-1668.

[30] McAllister, I. and T. Makkai (1991), "Correcting for the underreporting of drug use in opinion surveys", *International Journal of the Addictions*, 26, 945-961.

[31] McCaffrey, D.F., R.L. Pacula, B. Han and P. Ellickson (2008), "Marijuana Use and High School Dropout: The Influence of Unobservables", *NBER Working Paper*, no. 14102

[32] Mensch, B.S. and D.B. Kandel (1988), "Underreporting of substance abuse in a national longitudinal youth cohort", *Public Opinion Quarterly*, 52, 100-124.

[33] Mullahy, J. and J.L. Sindelar (1991), "Gender Differences in Labor Market Effects of Alcoholism", *American Economic Review Papers and Proceedings*, 81(2), 161-165.

[34] Osborn, A.F., N.R. Butler and A.C. Morris (1984), *The social life of Britain's five-year-olds*, London: Routledge.

[35] Pacula, R.L. (1997), "Women and Substance Use: Are Women Less Susceptible to Addiction?", *American Economic Review Papers and Proceedings*, 87(2), 454-459.

[36] Pacula, R.L. (1998), "Adolescent alcohol and marijuana consumption: is there really a gateway effect?", *NBER Working Paper*, no.6348.

[37] Pacula, R.L., J. Ringel and K.E. Ross (2003), "Does Marijuana Use Impair Human Capital Formation?", *NBER Working Paper*, no.9963.

[38] Pudney, S.E. (2004), "Keeping off the grass? An econometric model of cannabis consumption in Britain", *Journal of Applied Econometrics*, 19, 435-453.

[39] Pudney, S.E. (2007), "Rarely Pure and Never Simple: Extracting the Truth from Self-Reported Data on Substance Use", *Cemmap Working Paper 11*.

[40] Register, C.A. and D.R. Williams (1992), "Labor Market Effects of Marijuana and Cocaine Use among Young Men", *Industrial and Labor Relations Review*, 45(3), 435-448.

[41] Ringel, J.S., P.L. Ellickson and R.L. Collins (2006), "The Relationship between High School Marijuana Use and Annual Earnings among Young Adult Males", *Conteporary Economic Policy*, 24(1), 52-63.

[42] Rutter, M., J. Tizard and K. Whitmore (1970), *Education, Health and Behaviour*, London: Longman.

[43] Sander, W. (1998), "The Effects of Schooling and Cognitive Ability on Smoking and Marijuana Use by Young Adults", *Economics of Education Review*, 17(3), 317-324.

[44] Solowij, J. (1998), *Cannabis and Cognitive Functioning*, Cambridge: Cambridge University Press.

[45] Van Ours, J.C. (2007), "The Effects of Cannabis Use on Wages of Prime-age Males", *Oxford Bulletin of Economics and Statistics*, 69(5), 619-634.

[46] Van Ours, J.C. and J. Williams (2007), "Cannabis prices and dynamics of cannabis use", *Journal of Health Economics*, 26, 578-596.

Appendix 1 Data

Table A1 Cognitive tests used in the BCS

Test	Orig. range	Rel.	Mean[1]	Std.Dev.	n
Tests at age 5					
HFDT: Human Figure Drawing Test[2]	0-22(23)[3]	0.94[4]	0.463	0.116	12,824
CDT: Copying Designs Test	0-8	0.70[5]	0.591	0.247	13,028
EPVT: English Picture Vocabulary Test	0-56	0.96[6]	0.630	0.192	12,234
PT: Profile Test	0-16		0.432	0.249	12,451
Tests at age 10					
PLCT: Pictorial Language Comprehension Test	0-71		0.568	0.118	12,903
FMT: Friendly Math Test	0-72	0.92[7]	0.607	0.170	11,633
SERT: Shortened Edinburgh Reading Test	0-75	0.87[8]	0.603	0.207	11,641
BAS: British Ability Scales					
BAS-RD: Recall of Digits	0-34		0.659	0.126	11,512
BAS-M: Matrices	0-28		0.548	0.193	11,496
BAS-WD: Word Definitions	0-37		0.274	0.135	11,525
BAS-S: Similarities	0-21		0.575	0.123	11,473
Tests at age 16					
AT: Arithmetic Test	0-60		0.614	0.197	3,563
VT: Vocabulary Test	0-75		0.568	0.171	5,741
ST: Spelling Test	0-200		0.813	0.141	5,633

[1] The tests are normalised as proportion of the maximum possible score.

[2] Goodenough [1926] and Harris [1963]

[3] Two different scoring methods were used: according to the Harris method, the maximum score was 23, while it was 22 according to the Koppitz scoring method.

[4] Osborn et al. [1984]

[5] Osborn et al. [1984]

[6] Osborn et al. [1984]

[7] Butler et al. [1997]

[8] Butler et al. [1997]

Table A2 Fragmented sequences of cannabis use

Age 16	Age 30	Age 34	*n*	%
	Non-user	Non-user	2,073	47.12
	Past user	Non-user	1,159	26.35
	User	Occasional user	317	7.21
	Past user	Occasional user	317	7.21
	User	Regular user	209	4.75
	User	Non-user	201	4.57
	Non-user	Occasional user	89	2.02
	Past user	Regular user	28	0.64
	Non-user	Regular user	6	0.14
			4,399	100.00
Non-user	Non-user		408	47.39
Non-user	Past user		246	28.57
Non-user	User		141	16.38
User	Past user		30	3.48
User	User		30	3.48
User	Non-user		6	0.70
			861	100.00
Non-user		Non-user	166	51.71
Non-user		Occasional user	116	36.14
Non-user		Regular user	15	4.67
User		Occasional user	14	4.36
User		Regular user	7	2.18
User		Non-user	3	0.93
			321	100.00

Table A3 Cognitive ability by cannabis use and gender

Test	Age 16		Age 30			Age 34		
	Never used	Used	Never used	Used not last 12m	Used last 12m	Non user	Occas. user	Regular user
Age 5								
CDT[M]	0.627	0.641	0.577	0.636	0.616	0.611	0.634	0.583
CDT[F]	0.603	0.699	0.573	0.638	0.657	0.600	0.646	0.599
HFDT[M]	0.455	0.490	0.439	0.462	0.461	0.452	0.461	0.452
HFDT[F]	0.484	0.520	0.473	0.494	0.517	0.482	0.505	0.489
EPVT[M]	0.677	0.747	0.636	0.685	0.682	0.665	0.692	0.655
EPVT[F]	0.629	0.682	0.606	0.644	0.648	0.623	0.654	0.619
PT[M]	0.439	0.475	0.423	0.446	0.456	0.437	0.454	0.448
PT[F]	0.431	0.464	0.417	0.447	0.461	0.425	0.453	0.443
Age 10								
PLCT[M]	0.595	0.605	0.572	0.598	0.596	0.590	0.603	0.575
PLCT[F]	0.570	0.595	0.553	0.581	0.586	0.566	0.588	0.578
FMT[M]	0.655	0.697	0.606	0.648	0.649	0.636	0.661	0.615
FMT[F]	0.623	0.685	0.595	0.630	0.654	0.615	0.643	0.639
SERT[M]	0.635	0.710	0.570	0.629	0.636	0.610	0.646	0.586
SERT[F]	0.652	0.721	0.608	0.661	0.697	0.636	0.676	0.686
BAST-M[M]	0.579	0.621	0.528	0.570	0.579	0.556	0.589	0.531
BAST-M[F]	0.582	0.641	0.553	0.595	0.614	0.576	0.609	0.607
BAST-S[M]	0.604	0.654	0.571	0.605	0.614	0.591	0.619	0.586
BAST-S[F]	0.581	0.621	0.558	0.589	0.606	0.572	0.602	0.610
BAST-WD[M]	0.310	0.367	0.273	0.310	0.319	0.294	0.326	0.288
BAST-WD[F]	0.275	0.323	0.249	0.289	0.310	0.267	0.297	0.298
BAST-RD[M]	0.666	0.685	0.649	0.663	0.669	0.659	0.673	0.650
BAST-RD[F]	0.667	0.702	0.652	0.683	0.695	0.663	0.685	0.696
Age 16								
VOC[M]	0.589	0.609	0.547	0.581	0.591	0.570	0.599	0.560
VOC[F]	0.586	0.605	0.562	0.593	0.611	0.577	0.609	0.598
SPEL[M]	0.801	0.801	0.784	0.803	0.792	0.797	0.808	0.765
SPEL[F]	0.833	0.809	0.827	0.837	0.844	0.832	0.838	0.853
MATH[M]	0.648	0.664	0.600	0.647	0.622	0.636	0.638	0.566
MATH[F]	0.631	0.637	0.608	0.632	0.649	0.623	0.635	0.623

"Used at 16" groups the following four categories: *Yes, not in the past year; Yes, once in the past year; Yes, 2-9 times in the past year; Yes, 10+ times in the past year.*

"Non user at 34": note the question refers to the time period since last interview.

"Occasional user at 34" groups the following four categories: *Two or three times a month; Once a month; Less often/only on special occasions; I never use cannabis nowadays.*

"Regular user at 34" groups the following three categories: *On most days; Two or three times a week; Once a week.*

Table A4 Cannabis use by gender

	Males		*Females*	
Age 16	*n*	%	*n*	%
Never used	2,867	92.63	3,321	93.76
Used	228	7.37	221	6.24
	3,095	100.00	3,542	100.00
Age 30	*n*	%	*n*	%
Never used	2,135	39.75	3,330	58.40
Used, not in last 12 months	2,006	37.35	1,720	30.16
Used in last 12 months	1,230	22.90	652	11.43
	5,371	100.00	5,702	100.00
Age 34	*n*	%	*n*	%
Non-user	3,185	69.30	4,161	83.30
Occasional user	1,036	22.54	699	13.99
Regular user	375	8.16	135	2.70
	4,596	100.00	4,995	100.00

Table A5 Completed sequences of cannabis use by gender

	Males		*Females*	
Type	*n*	%	*n*	%
Abstainers	719	37.47	1,418	56.18
Middle Experimenters	521	43.42	550	49.73
Middle Recreational Users	173	14.42	118	10.67
Light Relapsers I	152	12.67	108	9.76
Quitters I	75	6.25	79	7.14
Middle Regular Users	100	8.33	31	2.80
Early Experimenters	37	3.08	82	7.41
Late Starters I	40	3.33	48	4.34
Early Recreational Users	30	2.50	26	2.35
Early Regular Users	25	2.08	14	1.27
Light Relapsers II	11	0.92	18	1.63
Heavy Relapsers I	14	1.17	9	0.81
Quitters II	12	1.00	7	0.63
Recanters	6	0.50	12	1.08
Heavy Relapsers II	3	0.25	3	0.27
Late Starters II	1	0.08	1	0.09
	1,200	100.00	1,106	100.00

Table A6 Synthetic sequences of cannabis use and cognitive ability

Test	Abstainers	Experimenters	Recreational users	Regular users
Age 5				
CDT	0.601	0.662	0.648	0.599
HFDT	0.468	0.486	0.486	0.474
EPVT	0.641	0.681	0.697	0.668
PT	0.426	0.444	0.471	0.454
Age 10				
PLCT	0.570	0.602	0.608	0.590
FMT	0.629	0.670	0.676	0.633
SERT	0.633	0.686	0.689	0.654
BAST-M	0.575	0.618	0.624	0.590
BAST-S	0.579	0.616	0.626	0.608
BAST-WD	0.274	0.319	0.329	0.305
BAST-RD	0.657	0.684	0.690	0.662
Age 16				
VOC	0.579	0.606	0.616	0.593
SPEL	0.820	0.833	0.824	0.805
MATH	0.634	0.672	0.665	0.587

The "experimenters" group the categories "early experimenters", "middle experimenters", "quitters I" and "quitters II"; the "recreational users" group the categories "late starters I", "light relapsers I", "middle recreational users", "light relapsers II" and "early recreational users"; the "regular users" group the categories "late starters II", "heavy relapsers I", "middle regular users", "heavy relapsers II" and "early regular users".

Table A7 Synthetic sequences of cannabis use and ability by gender

Test	Abstainers	Experim.	Recreational users	Regular users
Age 5				
CDT[M]	0.614	0.668	0.659	0.611
CDT[F]	0.593	0.657	0.632	0.573
HFDT[M]	0.448	0.467	0.469	0.467
HFDT[F]	0.479	0.502	0.507	0.492
EPVT[M]	0.672	0.709	0.716	0.683
EPVT[F]	0.625	0.655	0.671	0.633
PT[M]	0.440	0.447	0.465	0.470
PT[F]	0.418	0.441	0.478	0.417
Age 10				
PLCT[M]	0.581	0.613	0.620	0.589
PLCT[F]	0.564	0.592	0.592	0.593
FMT[M]	0.644	0.690	0.699	0.633
FMT[F]	0.621	0.651	0.647	0.633
SERT[M]	0.612	0.681	0.692	0.641
SERT[F]	0.644	0.691	0.685	0.686
BAST-M[M]	0.566	0.616	0.631	0.579
BAST-M[F]	0.580	0.619	0.616	0.616
BAST-S[M]	0.585	0.630	0.644	0.607
BAST-S[F]	0.575	0.603	0.605	0.612
BAST-WD[M]	0.286	0.338	0.354	0.305
BAST-WD[F]	0.268	0.303	0.299	0.307
BAST-RD[M]	0.656	0.680	0.683	0.648
BAST-RD[F]	0.658	0.687	0.697	0.693
Age 16				
VOC[M]	0.571	0.606	0.610	0.589
VOC[F]	0.582	0.606	0.624	0.600
SPEL[M]	0.790	0.818	0.811	0.781
SPEL[F]	0.833	0.843	0.839	0.844
MATH[M]	0.634	0.696	0.662	0.587
MATH[F]	0.634	0.653	0.669	0.587

The “experimenters” group the categories “early experimenters”, “middle experimenters”, “quitters I” and “quitters II”; the “recreational users” group the categories “late starters I”, “light relapsers I”, “middle recreational users”, “light relapsers II” and “early recreational users”; the “regular users” group the categories “late starters II”, “heavy relapsers I”, “middle regular users”, “heavy relapsers II” and “early regular users”.

Table A8 Summary Statistics: means of covariates by sequence type

Variable	Abstainers	Experim.	Recreational users	Regular users
Male	0.335	0.479	0.560	0.719
Mother's age at birth/100	0.264	0.261	0.267	0.265
Mother's education beyond MSLA	0.382	0.448	0.459	0.406
Father's education beyond MSLA	0.367	0.438	0.466	0.367
No husband at birth	0.030	0.039	0.036	0.070
Mother's husband SCI at birth	0.051	0.083	0.086	0.039
Mother's husband SCII at birth	0.120	0.154	0.168	0.094
Mother's husband SCIIINM at birth	0.149	0.145	0.134	0.156
Mother's husband SCIIIM at birth	0.472	0.412	0.428	0.367
Parity/10	0.096	0.098	0.104	0.137
Parental divorce 10-16	0.069	0.095	0.076	0.141
Father illness or disability 10-16	0.130	0.139	0.161	0.148
Accident 10-16	0.336	0.380	0.470	0.5
Younger siblings at 16	0.635	0.702	0.616	0.617
Dummy for missing questionnaire at 16	0.029	0.033	0.027	0.008
Father job loss 10-16	0.032	0.049	0.032	0.039
Dummy for missing father job loss	0.386	0.432	0.461	0.562
Religion: Christian	0.110	0.068	0.055	0.062
Religion: Other	0.125	0.151	0.147	0.148
Dummy for missing religion	0.293	0.355	0.359	0.437
Seizures	0.376	0.413	0.413	0.420
n	1,603	1,035	523	128

The sample includes both males and females.

Seizures has been constructed as the number of all controlled drugs seized in 1986 by Police Force Area and the mid-year population of the PFA.

Legend: MSLA=Minimum School Leaving Age; SC=Social Class.

Table A9 Estimated Coefficients: sequence analysis

Variable	Experim.	Recreational users	Regular users
Ability	0.395	0.373	0.257
	(0.061)	(0.076)	(0.130)
Male	0.567	0.846	1.548
	(0.088)	(0.111)	(0.234)
Mother's age at birth/100	-3.368	-1.041	-3.679
	(1.079)	(1.294)	(2.384)
Mother's education beyond MSLA	0.091	0.092	0.216
	(0.096)	(0.125)	(0.226)
Father's education beyond MSLA	0.098	0.230	0.090
	(0.105)	(0.132)	(0.243)
No husband at birth	0.346	0.405	0.542
	(0.243)	(0.314)	(0.450)
Mother's husband SCI at birth	0.319	0.331	-0.715
	(0.214)	(0.263)	(0.582)
Mother's husband SCII at birth	0.095	0.192	-0.701
	(0.166)	(0.210)	(0.387)
Mother's husband SCIIINM at birth	-0.146	-0.227	-0.415
	(0.162)	(0.211)	(0.333)
Mother's husband SCIIIM at birth	-0.099	0.028	-0.623
	(0.122)	(0.157)	(0.254)
Parity/10	1.508	1.267	3.233
	(0.452)	(0.578)	(0.831)
Parental divorce 10-16	0.412	0.232	0.779
	(0.156)	(0.205)	(0.310)
Father illness or disability 10-16	0.125	0.248	0.251
	(0.123)	(0.152)	(0.295)
Accident 10-16	0.099	0.439	0.394
	(0.089)	(0.110)	(0.210)
Younger siblings at 16	0.117	0.045	0.030
	(0.056)	(0.076)	(0.142)
Dummy for missing questionnaire at 16	0.188	0.075	-1.189
	(0.257)	(0.354)	(1.131)
Father job loss 10-16	0.697	0.262	0.481
	(0.220)	(0.322)	(0.531)
Dummy for missing father job loss	0.031	0.165	0.351
	(0.117)	(0.152)	(0.273)
Religion: Christian	-0.527	-0.748	-0.325
	(0.162)	(0.217)	(0.454)
Religion: Other	0.229	0.228	0.328
	(0.129)	(0.164)	(0.312)
Dummy for missing religion	0.215	0.048	0.126
	(0.127)	(0.165)	(0.303)
Seizures	0.729	0.774	0.798
	(0.181)	(0.231)	(0.447)
Constant	-1.052	-2.641	-3.598
	(0.299)	(0.371)	(0.649)
n	1,035	523	128

Table A9 (ctd.)

Cognitive Ability Production Function			
Male			0.104
			(0.046)
Mother's age/100			3.639
			(0.539)
Mother's education > MSLA			0.174
			(0.053)
Father's education > MSLA			0.295
			(0.056)
No husband			0.059
			(0.123)
Mother's husband SCI			0.559
			(0.122)
Mother's husband SCII			0.554
			(0.087)
Mother's husband SCIIINM			0.361
			(0.083)
Mother's husband SCIIIM			0.175
			(0.061)
Parity/10			-1.568
			(0.211)
Measurement Model			
ϕ_1	ϕ_2	ϕ_3	ϕ_4
0.440	0.410	0.562	0.336
(0.022)	(0.008)	(0.012)	(0.014)
α_1	α_2	α_3	α_4
0.1519	0.054	0.083	0.080
(0.006)	(0.002)	(0.004)	(0.006)
$\ln\sigma_1$	$\ln\sigma_2$	$\ln\sigma_3$	$\ln\sigma_4$
-1.711	-2.342	-1.879	-1.451
(0.023)	(0.014)	(0.014)	(0.020)
log-likelihood=694.19304			

Standard errors in parentheses.
The sample includes both males and females.
Legend: MSLA=Minimum School Leaving Age;
SC=Social Class.

Table A10 Summary Statistics: means of covariates by drug use

Variable	Not used by 16	Used by 16	Not used by 30	Used by 30	Not exp. by 30	Exper. by 30
Mother's age at birth/100	0.261	0.269	0.262	0.260	0.263	0.258
Mother's education > MSLA	0.414	0.550	0.346	0.404	0.360	0.411
Father's education > MSLA	0.402	0.533	0.337	0.395	0.361	0.385
No husband at birth	0.035	0.067	0.040	0.041	0.048	0.029
Mother's husband SCI	0.068	0.117	0.042	0.075	0.056	0.069
Mother's husband SCII	0.116	0.150	0.107	0.111	0.106	0.115
Mother's husband SCIIINM	0.155	0.167	0.151	0.156	0.146	0.166
Mother's husband SCIIIM	0.447	0.367	0.471	0.434	0.460	0.433
Parity/10	0.094	0.130	0.101	0.103	0.104	0.100
Parental divorce 10-16	0.079	0.117	0.064	0.082	0.065	0.090
Father illness or disability 10-16	0.126	0.217	0.107	0.111	0.116	0.098
Accident 10-16	0.454	0.417	0.357	0.421	0.384	0.411
Younger siblings at 16	0.665	0.517	0.578	0.575	0.560	0.601
Dummy for missing quest. at 16	0.032	0.050	0.150	0.157	0.157	0.148
Father job loss 10-16	0.031	0.017	0.021	0.031	0.024	0.030
Dummy for missing father job loss	0.650	0.490	0.525	0.587	0.548	0.581
Religion: Christian	(-)	(-)	0.080	0.048	0.068	0.050
Religion: Other	0.189	0.117	0.096	0.110	0.099	0.111
Dummy for missing religion	0.415	0.483	0.480	0.516	0.489	0.519
Seizures	0.393	0.495	0.380	0.410	0.389	0.410
O-levels grade a-c/10	0.294	0.328	0.240	0.256	0.240	0.264
Stay-on at school	0.528	0.500	0.431	0.465	0.438	0.471
Married 26-30	0.419	0.367	0.481	0.380	0.406	0.499
Change in activity	0.103	0.167	0.085	0.139	0.109	0.127
Change in marital status	0.748	0.767	0.779	0.729	0.755	0.743
Age 33 at interview	0.135	0.117	0.135	0.122	0.125	0.131
Promotion	0.355	0.400	0.347	0.329	0.322	0.360
New qualification	0.097	0.083	0.082	0.096	0.083	0.102
Part-time	0.008	0.033	0.012	0.015	0.016	0.011
Log-wage	2.213	2.295	2.182	2.225	2.182	2.247
n	1,138	60	815	1,130	1,190	755

The sample includes only males.

The category "Religion: Christian" has been grouped in the "Religion: Other" category for the analysis of drug use at 16, due to small cell size.

Legend: MSLA=Minimum School Leaving Age; SC=Social Class.

Table A11 Cognitive Ability and Drug Use

	Cannabis at 16	*Cannabis at 30 ever used*	*Cannabis at 30 experimented*
Ability	0.800	1.386***	1.046***
	(0.56)	(0.28)	(0.27)
Mother's age at birth/100	-1.693	-1.392*	-1.899***
	(1.580)	(0.730)	(0.74)
Mother's education beyond MSLA	0.218	0.087	0.089
	(0.150)	(0.070)	(0.070)
Father's education beyond MSLA	0.0467	0.001	-0.083
	(0.160)	(0.075)	(0.075)
No husband at birth	0.283	0.0280	-0.328*
	(0.350)	(0.160)	(0.170)
Mother's husband SCI at birth	0.214	0.243	0.021
	(0.300)	(0.160)	(0.150)
Mother's husband SCII at birth	0.150	-0.0446	-0.014
	(0.260)	(0.120)	(0.120)
Mother's husband SCIIINM at birth	0.115	-0.0152	0.048
	(0.250)	(0.110)	(0.110)
Mother's husband SCIIIM at birth	-0.009	-0.068	-0.063
	(0.200)	(0.082)	(0.082)
Parity/10	1.258**	0.569*	0.465
	(0.61)	(0.30)	(0.30)
Parental divorce 10-16	0.213	0.199*	0.220*
	(0.23)	(0.12)	(0.11)
Father illness or disability 10-16	0.429**	0.0523	-0.099
	(0.17)	(0.097)	(0.097)
Accident 10-16	-0.0577	0.200***	0.069
	(0.13)	(0.064)	(0.064)
Younger siblings at 16	-0.082	0.0241	0.022
	(0.094)	(0.042)	(0.042)
Dummy for missing questionnaire at 16	0.159	0.143	-0.008
	(0.34)	(0.094)	(0.094)
Father job loss 10-16	-0.161	0.431**	0.252
	(0.48)	(0.19)	(0.19)
Dummy for missing father job loss	0.396**	0.212**	0.0911
	(0.180)	(0.089)	(0.089)
Religion: Christian	-	-0.386***	-0.224*
	(-)	(0.130)	(0.13)
Religion: Other	-0.326	0.108	0.111
	(0.210)	(0.110)	(0.100)
Dummy for missing religion	-0.250	-0.032	0.059
	(0.180)	(0.093)	(0.093)
Seizures	0.731***	0.269**	0.180
	(0.240)	(0.120)	(0.120)
O-levels grade a-c	-	-0.110	0.027
	(-)	(0.110)	(0.100)
Stay-on at school	-	0.055	0.037
	(-)	(0.073)	(0.073)
Constant	-2.097***	-0.094	-0.235

Table A11 (ctd.)

Cognitive Ability Production Function			
Mother's age/100	0.614***	0.560***	0.563***
	(0.140)	(0.110)	(0.110)
Mother's education > MSLA	0.0225*	0.0262**	0.0264**
	(0.014)	(0.011)	(0.011)
Father's education > MSLA	0.0497***	0.057***	0.0568***
	(0.015)	(0.011)	(0.011)
No husband	-0.0488	-0.036	-0.035
	(0.034)	(0.025)	(0.025)
Mother's husband SCI	0.091***	0.105***	0.104***
	(0.028)	(0.023)	(0.023)
Mother's husband SCII	0.079***	0.077***	0.077***
	(0.024)	(0.019)	(0.019)
Mother's husband SCIIINM	0.055**	0.043**	0.043**
	(0.022)	(0.017)	(0.017)
Mother's husband SCIIIM	0.031*	0.032**	0.031**
	(0.017)	(0.013)	(0.013)
Parity/10	-0.243***	-0.239***	-0.240***
	(0.060)	(0.045)	(0.045)
Measurement Model			
ϕ_1	0.440***	0.436***	0.436***
	(0.035)	(0.027)	(0.027)
ϕ_2	0.382***	0.384***	0.384***
	(0.014)	(0.011)	(0.011)
ϕ_3	0.595***	0.583***	0.584***
	(0.020)	(0.016)	(0.016)
ϕ_4	0.345***	0.354***	0.354***
	(0.021)	(0.015)	(0.015)
α_1	1	1	1
	(-)	(-)	(-)
α_2	0.386***	0.387***	0.387***
	(0.032)	(0.024)	(0.025)
α_3	0.491***	0.526***	0.520***
	(0.049)	(0.039)	(0.039)
α_4	0.460***	0.442***	0.443***
	(0.063)	(0.048)	(0.049)
$\ln\sigma_1$	-1.692***	-1.688***	-1.690***
	(0.039)	(0.029)	(0.029)
$\ln\sigma_2$	-2.327***	-2.363***	-2.364***
	(0.028)	(0.022)	(0.022)
$\ln\sigma_3$	-1.896***	-1.876***	-1.873***
	(0.026)	(0.021)	(0.021)
$\ln\sigma_4$	-1.407***	-1.415***	-1.416***
	(0.022)	(0.017)	(0.017)
$\ln\sigma_u$	0.153***	0.152***	0.153***
	(0.009)	(0.007)	(0.007)

Standard errors in parentheses. n=1,198.

The sample includes only males.

Legend: MSLA=Minimum School Leaving Age; SC=Social Class.

Table A12 Cognitive Ability, Drug Use and Wages

Dependent Variable: log hourly net wage at 34		
Ever used cannabis by 16	0.216	0.0526
	(0.170)	(0.057)
Ability	-	0.374***
	-	(0.110)
Mother's education > MSLA	0.0408	0.0356
	(0.028)	(0.028)
Father's education > MSLA	0.122***	0.104***
	(0.029)	(0.029)
O-levels grade a-c	0.231***	0.196***
	(0.041)	(0.042)
Stay-on at school	0.114***	0.099***
	(0.030)	(0.030)
Married 26-30	0.111***	0.106***
	(0.027)	(0.027)
Change in activity	-0.090**	-0.088**
	(0.042)	(0.042)
Change in marital status	-0.053*	-0.051*
	(0.030)	(0.030)
Age 33 at interview	-0.013	-0.014
	(0.036)	(0.036)
Promotion	0.095***	0.097***
	(0.027)	(0.026)
New qualification	0.0105	0.001
	(0.042)	(0.042)
Part-time	-0.166	-0.158
	(0.130)	(0.130)
Constant	1.983***	1.942***
	(0.032)	(0.036)
$\ln(\sigma_w)$	-0.849***	-0.861***
	(0.021)	(0.021)
Correlation between residuals	-0.171	
in drug use and wage equations	(0.183)	

Column 2: Model without cognitive ability.

Column 3: Model with cognitive ability.

Table A12 (ctd.)

Drug Use - Ever used cannabis by 16		
Ability	-	0.799
	-	(0.570)
Mother's age at birth/100	-1.116	-1.706
	(1.530)	(1.580)
Mother's education beyond MSLA	0.219	0.218
	(0.150)	(0.150)
Father's education beyond MSLA	0.082	0.0466
	(0.160)	(0.160)
No husband at birth	0.184	0.286
	(0.350)	(0.350)
Mother's husband SCI at birth	0.319	0.213
	(0.290)	(0.300)
Mother's husband SCII at birth	0.234	0.150
	(0.250)	(0.260)
Mother's husband SCIIINM at birth	0.156	0.115
	(0.240)	(0.250)
Mother's husband SCIIIM at birth	0.030	-0.009
	(0.200)	(0.200)
Parity/10	1.090*	1.260**
	(0.590)	(0.610)
Parental divorce 10-16	0.188	0.214
	(0.220)	(0.230)
Father illness or disability 10-16	0.416**	0.429**
	(0.170)	(0.170)
Accident 10-16	-0.070	-0.058
	(0.130)	(0.130)
Younger siblings at 16	-0.082	-0.082
	(0.093)	(0.094)
Dummy for missing questionnaire at 16	0.160	0.157
	(0.350)	(0.340)
Father job loss 10-16	-0.151	-0.160
	(0.460)	(0.480)
Dummy for missing father job loss	0.386**	0.397**
	(0.180)	(0.180)
Religion: Other	-0.359*	-0.325
	(0.210)	(0.210)
Dummy for missing religion	-0.248	-0.251
	(0.180)	(0.180)
Seizures	0.761***	0.729***
	(0.240)	(0.240)
Constant	-2.101***	-2.095***
	(0.440)	(0.450)

Column 2/3: Model without/with cognitive ability.

Table A12 (ctd.)

Cognitive Ability Production Function	
Mother's age/100	0.626***
	(0.14)
Mother's education > MSLA	0.0235*
	(0.014)
Father's education > MSLA	0.0491***
	(0.015)
No husband	-0.0530
	(0.034)
Mother's husband SCI	0.092***
	(0.028)
Mother's husband SCII	0.079***
	(0.024)
Mother's husband SCIIINM	0.056**
	(0.022)
Mother's husband SCIIIM	0.031*
	(0.017)
Parity/10	-0.244***
	(0.060)
Measurement Model	
ϕ_1	0.436***
	(0.035)
ϕ_2	0.381***
	(0.014)
ϕ_3	0.593***
	(0.020)
ϕ_4	0.346***
	(0.021)
α_1	1
	(-)
α_2	0.382***
	(0.031)
α_3	0.495***
	(0.049)
α_4	0.450***
	(0.063)
$\ln\sigma_1$	-1.693***
	(0.039)
$\ln\sigma_2$	-2.323***
	(0.028)
$\ln\sigma_3$	-1.899***
	(0.027)
$\ln\sigma_4$	-1.405***
	(0.022)
$\ln\sigma_u$	0.153***
	(0.009)

Standard errors in parentheses. n=1,198.

The sample includes only males.

Legend: MSLA=Minimum School Leaving Age; SC=Social Class.

Table A13 Cognitive Ability, Drug Use and Wages

Dependent Variable: log hourly net wage at 34		
Ever used cannabis by 30	0.603***	0.0295
	(0.049)	(0.020)
Ability	-	0.348***
	-	(0.086)
Mother's education > MSLA	-0.002	0.0122
	(0.027)	(0.023)
Father's education > MSLA	0.101***	0.103***
	(0.027)	(0.023)
O-levels grade a-c	0.258***	0.221***
	(0.040)	(0.035)
Stay-on at school	0.0695**	0.0658***
	(0.028)	(0.024)
Married 26-30	0.118***	0.112***
	(0.021)	(0.021)
Change in activity	-0.138***	-0.128***
	(0.032)	(0.032)
Change in marital status	-0.0483**	-0.0487**
	(0.024)	(0.024)
Age 33 at interview	-0.0221	-0.0212
	(0.029)	(0.029)
Promotion	0.0943***	0.0955***
	(0.021)	(0.021)
New qualification	-0.006	-0.0219
	(0.034)	(0.034)
Part-time	-0.189**	-0.180**
	(0.084)	(0.086)
Constant	1.701***	1.975***
	(0.037)	(0.029)
$\ln(\sigma_w)$	-0.676***	-0.855***
	(0.030)	(0.016)
Correlation between residuals in drug use and wage equations	-0.698*** (0.039)	

Column 2: Model without cognitive ability.

Column 3: Model with cognitive ability.

Table A13 (ctd.)

Drug Use - Ever used cannabis by 30		
Ability	-	1.393***
	-	(0.28)
Mother's age at birth/100	0.537	-1.423*
	(0.59)	(0.73)
Mother's education beyond MSLA	0.0909	0.0870
	(0.068)	(0.070)
Father's education beyond MSLA	0.0514	0.001
	(0.071)	(0.075)
No husband at birth	-0.136	0.031
	(0.13)	(0.16)
Mother's husband SCI at birth	0.330***	0.240
	(0.13)	(0.16)
Mother's husband SCII at birth	0.0989	-0.0462
	(0.096)	(0.12)
Mother's husband SCIIINM at birth	0.0998	-0.0179
	(0.087)	(0.11)
Mother's husband SCIIIM at birth	0.0569	-0.0705
	(0.065)	(0.082)
Parity/10	-0.109	0.578*
	(0.23)	(0.31)
Parental divorce 10-16	0.149	0.198*
	(0.097)	(0.12)
Father illness or disability 10-16	0.0308	0.0520
	(0.078)	(0.097)
Accident 10-16	0.151***	0.200***
	(0.052)	(0.064)
Younger siblings at 16	0.0234	0.0237
	(0.034)	(0.042)
Dummy for missing questionnaire at 16	0.261***	0.140
	(0.076)	(0.094)
Father job loss 10-16	0.121	0.435**
	(0.16)	(0.19)
Dummy for missing father job loss	0.0564	0.214**
	(0.073)	(0.089)
Religion: Christian	-0.406***	-0.383***
	(0.11)	(0.13)
Religion: Other	0.064	0.108
	(0.086)	(0.11)
Dummy for missing religion	0.054	-0.034
	(0.076)	(0.093)
Seizures	0.385***	0.266**
	(0.098)	(0.12)
O-levels grade a-c	-0.0431	-0.111
	(0.10)	(0.11)
Stay-on at school	0.084	0.056
	(0.070)	(0.073)
Constant	-0.394**	-0.089
	(0.17)	(0.21)

Column 2/3: Model without/with cognitive ability.

Table A13 (ctd.)

Cognitive Ability Production Function	
Mother's age/100	0.572***
	(0.11)
Mother's education > MSLA	0.027**
	(0.011)
Father's education > MSLA	0.056***
	(0.011)
No husband	-0.0393
	(0.025)
Mother's husband SCI	0.107***
	(0.023)
Mother's husband SCII	0.078***
	(0.019)
Mother's husband SCIIINM	0.044***
	(0.017)
Mother's husband SCIIIM	0.033***
	(0.013)
Parity/10	-0.242***
	(0.044)
Measurement Model	
ϕ_1	0.432***
	(0.027)
ϕ_2	0.383***
	(0.011)
ϕ_3	0.581***
	(0.016)
ϕ_4	0.355***
	(0.015)
α_1	1
	(-)
α_2	0.384***
	(0.024)
α_3	0.530***
	(0.039)
α_4	0.432***
	(0.048)
$\ln\sigma_1$	-1.688***
	(0.029)
$\ln\sigma_2$	-2.360***
	(0.022)
$\ln\sigma_3$	-1.879***
	(0.021)
$\ln\sigma_4$	-1.413***
	(0.017)
$\ln\sigma_u$	0.152***
	(0.007)

Standard errors in parentheses. n=1,945.

The sample includes only males.

Legend: MSLA=Minimum School Leaving Age; SC=Social Class.

Table A14 Cognitive Ability, Drug Use and Wages

Dependent Variable: log hourly net wage at 34		
Experimented before 30	0.587***	0.0325
	(0.053)	(0.020)
Ability	-	0.349***
	-	(0.085)
Mother's education > MSLA	-0.002	0.0121
	(0.027)	(0.023)
Father's education > MSLA	0.127***	0.104***
	(0.027)	(0.023)
O-levels grade a-c	0.237***	0.220***
	(0.040)	(0.035)
Stay-on at school	0.070**	0.065***
	(0.028)	(0.024)
Married 26-30	0.107***	0.107***
	(0.021)	(0.021)
Change in activity	-0.139***	-0.126***
	(0.032)	(0.032)
Change in marital status	-0.049**	-0.0483**
	(0.024)	(0.024)
Age 33 at interview	-0.021	-0.0225
	(0.029)	(0.029)
Promotion	0.095***	0.0941***
	(0.021)	(0.021)
New qualification	0.001	-0.0222
	(0.034)	(0.034)
Part-time	-0.166**	-0.177**
	(0.084)	(0.086)
Constant	1.823***	1.981***
	(0.032)	(0.028)
$\ln(\sigma_w)$	-0.687***	-0.855***
	(0.031)	(0.016)
Correlation between residuals in drug use and wage equations	-0.678*** (0.043)	

Column 2: Model without cognitive ability.

Column 3: Model with cognitive ability.

Table A14 (ctd.)

Drug Use - Experimented before 30		
Ability	-	1.050***
	-	(0.270)
Mother's age at birth/100	0.019	-1.922***
	(0.600)	(0.740)
Mother's education beyond MSLA	0.105	0.089
	(0.068)	(0.070)
Father's education beyond MSLA	-0.051	-0.082
	(0.071)	(0.075)
No husband at birth	-0.436***	-0.325*
	(0.14)	(0.170)
Mother's husband SCI at birth	0.183	0.019
	(0.120)	(0.150)
Mother's husband SCII at birth	0.089	-0.014
	(0.098)	(0.120)
Mother's husband SCIIINM at birth	0.118	0.046
	(0.088)	(0.110)
Mother's husband SCIIIM at birth	0.0370	-0.064
	(0.067)	(0.082)
Parity/10	-0.0413	0.471
	(0.240)	(0.300)
Parental divorce 10-16	0.170*	0.219*
	(0.096)	(0.110)
Father illness or disability 10-16	-0.0711	-0.099
	(0.081)	(0.097)
Accident 10-16	0.054	0.0691
	(0.053)	(0.064)
Younger siblings at 16	0.0187	0.0221
	(0.034)	(0.041)
Dummy for missing questionnaire at 16	0.134*	-0.0103
	(0.078)	(0.094)
Father job loss 10-16	0.017	0.255
	(0.160)	(0.19)
Dummy for missing father job loss	-0.028	0.0927
	(0.074)	(0.089)
Religion: Christian	-0.313***	-0.221*
	(0.110)	(0.13)
Religion: Other	0.053	0.111
	(0.086)	(0.10)
Dummy for missing religion	0.113	0.0580
	(0.077)	(0.093)
Seizures	0.349***	0.178
	(0.099)	(0.12)
O-levels grade a-c	0.0406	0.0261
	(0.10)	(0.10)
Stay-on at school	0.0625	0.0368
	(0.070)	(0.073)
Constant	-0.595***	-0.232
	(0.170)	(0.21)

Column 2/3: Model without/with cognitive ability.

Table A14 (ctd.)

Cognitive Ability Production Function	
Mother's age/100	0.575***
	(0.11)
Mother's education > MSLA	0.0273**
	(0.011)
Father's education > MSLA	0.0559***
	(0.011)
No husband	-0.0380
	(0.025)
Mother's husband SCI	0.107***
	(0.023)
Mother's husband SCII	0.0770***
	(0.019)
Mother's husband SCIIINM	0.0443***
	(0.017)
Mother's husband SCIIIM	0.0330***
	(0.013)
Parity/10	-0.242***
	(0.044)
Measurement Model	
ϕ_1	0.432***
	(0.027)
ϕ_2	0.383***
	(0.011)
ϕ_3	0.581***
	(0.016)
ϕ_4	0.354***
	(0.015)
α_1	1
	(-)
α_2	0.384***
	(0.024)
α_3	0.525***
	(0.039)
α_4	0.433***
	(0.048)
$\ln\sigma_1$	-1.690***
	(0.029)
$\ln\sigma_2$	-2.361***
	(0.022)
$\ln\sigma_3$	-1.876***
	(0.021)
$\ln\sigma_4$	-1.414***
	(0.017)
$\ln\sigma_u$	-0.152***
	(0.007)

Standard errors in parentheses. n=1,945.

The sample includes only males.

Legend: MSLA=Minimum School Leaving Age; SC=Social Class.