Why Parents Worry: Initiation into Cannabis Use by Youth and their Educational Attainment

Jan C. van Ours* Jenny Williams†

October 26, 2006

Abstract

In this paper we investigate the extent to which initiation into cannabis use during youth affects educational attainment. In particular we focus on the relationship between the age at which initiation occurs and the probability of leaving formal education. We find that age of onset matters, with those initiating into cannabis use early in life having a much higher dropout rate than those who start later on. For boys the critical age for starting cannabis use is 17, and for girls it is 19. Initiation into cannabis use beyond the critical age has no adverse affect on educational attainment.

Keywords: cannabis use; age of initiation; educational attainment

JEL codes: C41, D12, I19

*Department of Economics, Tilburg University, CentER, and CEPR; email: vanours@uvt.nl.

†Department of Economics, University of Melbourne, Parkville, Australia; email: jenny.williams@unimelb.edu.au

1 Introduction

One of parents' greatest fears is that their child will become involved with drugs. Underlying this fear is the belief that drug use could lead to poor educational attainment, subsequent failure in the labor market, and without a good job to anchor their lives, an unhappy future. Viewed within a human capital framework, this scenario may find resonance. For example, drug use could lead teenagers to substitute time spent under the influence of drugs for time spent studying, resulting in poor academic achievement and an early exit from education. This is particularly a concern with cannabis because initiation into its use typically occurs during the teenage years, coinciding with the timing of critical decisions about investment in formal education, both at the extensive and intensive margins. There is, therefore, potential for youthful cannabis use to have a long lasting affect through its impact on the individual's stock of human capital. This paper investigates the extent to which this is the case by examining how the age of initiation into cannabis use effects subsequent educational attainment.

There is substantial evidence that early cannabis use is associated with lower levels of education (Macleod et al., 2004). What is less well understood is the extent to which this association reflects the causal impact of cannabis use on education outcomes. Associations will not reflect causal effects if, for example, those who self-select into cannabis use differ from those who do not use cannabis in ways that also impact on their academic achievement (selection on unobservables), or if poor educational attainment is both a cause and a consequence of youthful initiation into cannabis (reverse causality). The presence of either of these issues will render cannabis use endogenous to decisions regarding education. If unaccounted for, this endogeneity will lead to inconsistent estimates of the effect of cannabis use on education. Given the obvious objections to using an experimental approach, economics is particularly well placed to address these issues and hence obtain reliable estimates of the impact of youthful drug use on educational attainment.

Despite this, there are only a handful of relevant studies in the economics literature. These studies, reviewed in the following section, find that drug use during high-school reduces the number of years of education completed by between 0.2 and 1 year. None of the previous studies, however, investigate the role of the age at which initiation occurs, nor do they consider the educational consequences of drug use beyond the high-school years. Thus, there is no evidence on whether initiation into cannabis use at age 14 is more or less damaging in terms of educational outcomes than initiation at age 17, or whether there remains a negative effect of use beyond high-school. The aim of this paper is to provide answers to these questions. Knowledge about the relative impact of up-take at different ages is useful from a policy perspective because it can help in targeting strategies that aim to minimize the harm associated with cannabis use.

In this paper, we adopt two approaches to addressing the potential endogeneity of initiation into cannabis in order to estimate its impact on the decision to leave education. First, we use instrumental variables (IV) estimation. An advantage of this approach is that, since it is used in the previous studies, we are able to compare our results with those in the literature. The IV approach is, however, ill-suited to studying aspects related to the timing of transitions into cannabis use and out of education. These transitions are more naturally viewed within a duration framework. For this reason, our second approach uses a bivariate duration model to investigate the impact of the age at which initiation into cannabis use occurs on the probability of leaving formal education.

A benefit of using two estimation strategies is that we are able to examine the robustness of our results to the identifying assumptions employed. In the case of IV estimation, identification of the effect of initiation into cannabis use on educational attainment is based on exclusion restrictions, while the duration approach relies on the timing of events. Independent of the estimation strategy and hence identifying assumption employed, we find that initiation into cannabis use, particularly in the early teenage years, is associated with fewer years of education. Also consistent across estimation methods is the finding that, for females but not for males, the unobserved characteristics that lead individuals to initiate cannabis use also make them more likely to pursue more formal education. Consequently, failing to account for this correlation leads to an underestimate of the negative impact of initiation into cannabis use on female's educational attainment. Finally, we find that the age at which initiation occurs matters in terms of educational attainment and that its effect differs across gender. More specifically, we find that up-take of cannabis before the ages of 18 for males, and before the age of 20 for girls, leads to a reduction in their expected years of completed education. Moreover the magnitude of the effect is larger for females than males.

The rest of this paper is laid out as follows. Section 2 reviews economic studies that investigate the impact of early cannabis use on educational attainment. Section 3 describes the data used in this paper. Section 4 presents the IV framework and empirical results on the relationship between youthful initiation into cannabis use and educational success, while section 5 contains the econometric set-up and results for the duration analysis. Section 6 concludes with a discussion of our findings.

2 Literature Review

The focus of this review is on studies from the economics literature, and particularly those which empirically address the potential endogeneity of cannabis use in decisions about formal education. For a wider review of research on the relationship between drug use and education, see Chatterji (2006) or Pacula et. al (2005).

The economics literature on the relationship between youthful cannabis use and education is limited and relies solely on data from the US. The first two published studies focus on the association between drug use and completing high-school. Yamada, Kendix and Yamada (1996) report that heavy cannabis use in twelfth grade is associated with a reduced probability of graduating, while and Bray et al. (2000) find that initiation into cannabis use prior to age 16, 17 and 18 is associated with an increased probability of dropping out of high-school at these ages.

An issue not addressed in these early studies is whether the empirical relationship between drug use and educational attainment is causal.[1] Three studies have attempted to address this issue. First, Register, Williams and Grimes (2001) use data on males from the National Longitudinal Study of Youth to examine the impact of drug use by the age of 18 on the number of years of education completed. Using a two-step estimator to account for the potential endogeneity of drug use, they find that, on average, male adolescent drug use is associated with a reduction of around 1 year in educational attainment, where this result is driven by the whites in the sample. Unfortunately, the authors report none of the usual specification tests associated with instrumental variable estimation, making it is difficult to evaluate the reliability of their estimates. Information on the statistical merits of their results would be most useful in light of the fact that the impact of adolescent cannabis use is identified from variation in the respondents' religiosity, and it is not obvious that values associated with being religious are orthogonal to educational attainment.

Chatterji (2006) exploits the unusually rich National Education Longitudinal Study to examine the impact of past month cannabis use in 10^{th} and 12^{th} grade on subsequent years of education completed. In addition to employing an IV approach, Chatterji also attempts to mitigate the potential omitted variable bias associated with OLS estimation by controlling for (typically unobserved) preexisting individual factors that may confound the relationship between drug use in high-school and educational attainment. The results from doing so suggest that past month use of cannabis in the 10^{th} or 12^{th} grade reduces educational attainment by 0.2 years. Unfortunately, the instruments employed by Chatterji in the IV estimation are only weakly correlated with the measures of cannabis use.[2] Given the potential biases

[1]Bray et al. (2000) discuss (but do not present results from) efforts to address the potential endogeneity of initiation into drug use. No discussion of the variables used to identify the effect of drug use is provided.

[2]The F-statistics for testing the significance of the instruments in the first stage regressions range from 3.14 to 8.51 when drug use is measured by 10^{th} grade use, and 1.23 to 5.90 when drug use is measured by 12^{th} grade use.

associated with weak instruments, it is perhaps not surprising that the Hausman test fails to find a significant difference between the OLS and IV estimates in this paper. A further source of concern is that specifications that include pre-existing individual factors produce IV estimates of the impact of cannabis use on educational attainment that are incorrectly signed. This, along with the potential problem of reverse causality, casts some doubt on the reliability of the OLS estimates from specifications that include these variables.[3]

The relationship between the intensity and persistence of cannabis use and educational attainment is studied by Pacula et al. (2005). This study uses the RAND Adolescent Panel Survey, which follows individuals who were enrolled in thirty middle schools in California and Oregon from 7^{th} grade through to age of 23. Using an IV approach, the authors find no evidence that participation in cannabis use in the 7^{th} or 12^{th} grade impacts on educational attainment, as measured by highest level of education completed, or by graduation from high-school. Frequency of cannabis use in the 12^{th} grade, and persistent use (defined as using at least 3 times in the month prior to survey in the 10^{th} and 12^{th} grades) is, however, found to reduce educational attainment. For example, persistent use is estimated to reduce the educational attainment by 0.7 years. As with Chatterji (2006), weak instruments are an issue in this study.[4] This raises the question of the robustness of inference regarding the impact of frequent drug use.

These studies are unanimous in finding that cannabis use in high-school reduces educational attainment. This raises the question of whether use is more harmful in terms of education outcomes at some ages compared to others. For example, is

[3]Some of the variables used to control for pre-existing conditions, such as low grade 10 math score, low 8^{th} grade math score, repeated a grade, number of times parents were called in for problem behavior in the 8^{th} grade, are likely themselves to be endogenous. Since omitted variable bias is being traded for endogeneity bias, it is not clear that the OLS results provide a consistent estimator of the causal impact of drug use on educational attainment.

[4]The F-statistic on the joint significance of the instruments is reported to be 3.36 for the model for the frequency of cannabis use in grade 12, and 12.92 for participation in use in year 12.

the reduction in years of educational attainment greater for those who start using cannabis at earlier ages? Or is it up-take at critical ages, such as the age at which final exams for high-school graduation are taken, that is more harmful? The focus on cannabis use during high-school in previous studies also raises the question of whether there are any adverse effects of up-take after high-school in terms of educational attainment? Questions of a more methodological nature also arise from the literature. Specifically, as all of the studies that account for the endogeneity of youthful cannabis use rely on the same empirical methodology, instrumental variable estimation, they are all subject to the same frailties. This raises the question of whether their results are robust to alternative identifying assumptions and estimation strategies. Providing answers to these questions is the motivation for what follows.

3 Data

3.1 Australian National Drug Strategy Household Survey

This research draws on information collected in the 2001 Australian National Drug Strategy Household Survey (NDSHS). The NDSHS is designed to provide data on the extent of drug use by the non-institutionalized civilian population aged fourteen years and older in Australia. In addition to asking individuals whether they have ever used or currently use various licit and illicit drugs, the NDSHS also asks those who report having ever used each substance the age at which it was first used. This, along with information on the respondents' highest level of schooling and any post-school qualifications make these data useful for examining the impact of the age of initiation into cannabis use on education outcomes.

In the analysis that follows, we make use of two measures of the age of initiation into cannabis use and three measures of educational attainment. Age of initiation into cannabis use is measured by an indicator for first use at or before the age of 15 and by the actual age at which initiation occurred. The former is used in the

IV analysis while the latter is used in the duration analysis. The three measures of educational attainment are: (1) the age at which the respondent left school, (2) an indicator for graduated from high-school, and (3) an indicator for completed at least an undergraduate degree. All three measures are used in the IV estimation, whereas the outcome of interest for the duration analysis is the age at which the individual left school.

As we wish to ascertain the impact of age of initiation into cannabis use on the decision to leave formal education, we focus on those who can reasonably be considered to have completed their education. For this reason we limit our sample to individuals aged 25-50 years old and who do not report that their primary activity is study. We have 11,793 observations for which we have complete data on education, cannabis use, and the other control variables for this age group. Summary statistics for the outcomes and interest and other explanatory variables are reported in Table 1. They show that 9% of the sample initiated into cannabis use by the age of 15, and that amongst those who have ever used cannabis, the average age of initiation is 19 years. In terms of educational attainment, on average, sample members leave school at 18 years of age, with 55% having high-school graduation as their highest level of formal education, and 25% having an undergraduate or post-graduate degree as their highest level of education. In terms of demographic characteristics, 42% of the sample is male, 78% are Australian born and 26% live in rural locations. The average age of the sample is 37 years. The IV analysis also makes use of information on cigarette use. As seen form Table 1, 50% of the sample smoked cigarettes daily at some point in their life and a further 22% have smoked cigarettes, but never used them daily. In terms of initiation into cigarette use, 5% of the sample first smoked a full cigarette by the age of 10, 2% at the age of 11 and a further 7% at the age of 12.

As with previous research studying transitions in substance use using cross-sectional data, this study is subject to potential measurement error problems. First, using retrospective information about when individuals start cannabis use poses the potential problem of recall error. As discussed below, we find some evidence of recall error in the reported age of initiation into cannabis use for those who initiate into use after the age of 20. This is not surprising as we are dealing with respondents aged 20 to 50 years, and initiation into cannabis use typically occurs in the mid to late teens. To the extent that respondent's make errors in the age they report first using cannabis, our parameter estimates are likely to be biased towards zero. However, since initiation into use after the age of 25 is fairly rare, we do not anticipate large effects form this source of measurement error.

The education outcome 'age left education' may also suffer from measurement error. We construct the age at which an individual left formal education using information on schooling and post-school qualifications along with historic information on the structure of the education system in each state.[5] For example, if the highest qualification a person achieved is graduating from high school, they are attributed an age for completing education which is equal to the school starting age (five) plus the number of years of education required to complete high-school in their state of residence. If the highest qualification is an under-graduate degree, we add a further four years to their school-leaving age.[6] A masters degree is assumed to take an additional two years after the undergraduate degree, and a PhD a further three years (which is the length of government scholarships). To the extent that individuals repeat grades at school or take longer than the standard time to complete university study, we will tend to under-state the age at which they complete formal education. However, the other two indicators of educational attainment, graduating from high-

[5]Post school qualifications gained at institutions other than universities are generally vocational in nature and for this reason are not included in the age left school variable, which reflects formal education only. Differing education systems across states mean that students in Queensland are 17 when they graduate from high-school whereas their New South Wales counterparts are 18. The information on the school systems was generously provided by Chris Ryan from the Australian National University.

[6]A (typical) undergraduate degree takes three years to complete in Australia. We allow four years for completion to account for the longer duration for combined degrees, honors degrees, and to account for the fact that the progression rate from one year to the next is less than 1.

school and graduating from university, are not subject to these measurement issues. The indicator for having graduating from university is constructed from information reported on the highest post-school qualification and is equal to one if the highest post-school qualification is a bachelors degree, a masters degree or a Ph.D. The indicator for graduating from high-school is equal to one if the respondent has received a bachelor's degree or higher degree (masters or PhD) or if they report that their highest level of schooling is a year 12 certificate, which is the Australian equivalent of graduating from high-school.

3.2 Descriptive Statistics

Figures 1 and 2 give a graphical illustration of the relationship between age and initiation into cannabis use. Figure 1 graphs the probability of starting cannabis use at each age, conditional on not having been a user up to that age. The figure shows that initiation into cannabis use begins for the sample at age 10, with 0.1% of males and females reporting first use at that age. The first peak in the probability of up-take is at age 16, when 9.2% of males and 6.1% of females who had not previously used cannabis initiate use. The mean peak is at age 18, with 12.3% of males and 9.0% of females initiating use, but there is also a peak of 8.5% for males and 5.6% for females at age 20. Subsequent peaks are at age 25, 30, 35, and 40; these peaks in the age-specific starting probabilities point to bundling in the recollection of the starting age. Figure 1 also shows clearly that initiation into cannabis use rarely occurs beyond age 25. Figure 2 shows the cumulative probability of initiating cannabis use by age. At age 15, 9.4% of males and 8.2% of females in our sample have started cannabis use. This increases to 43.4% of males and 34.9% of females at age 20 and 50.2% of males and 42.2% of females at age 25. By the age of 50, 54.9% of males and 47.0% of females have used cannabis at some point in their life.

Figures 3 and 4 provide similar information on the relationship between age and the decision to leave formal education. As shown, the graphs for males and females are virtually the same. Figure 3 shows that individuals start leaving the formal education sector at age 13, although the probability of doing so at that age is very small. There is a clear peak in the probability of leaving education at age 18, which is the age for completing high-school in most Australian states. At age 19 and 20 none of the individuals in our sample are observed to leave education, but there are further peaks at ages 22 and 24. This pattern reflects the fact that the NDSHS measures the highest educational qualification attained and not the age the respondent left school per-se. Figure 4 shows the cumulative probability of leaving education.

The joint distribution of initiation into cannabis use and educational attainment are reported for males and females in Table 2. As shown, the marginal distribution for educational attainment is very similar for males and females but the marginal distribution of age of initiation into cannabis use is different. The null hypothesis that educational attainment and age of initiation into cannabis are independent is strongly rejected by the Pearson χ^2 test for both males and females, with a p-value of 0.000 for each gender.[7] Overall, 45% of males and females drop out of formal education before completing high-school, and 29% have a level of educational attainment equal to high-school graduate. Males have a slightly lower probability of an undergraduate degree being their highest level of education (15%) compared to females (17%), and a slightly higher probability of having a postgraduate degree (10%) compared to females (9%). Moreover these proportions are statistically different at the 5% level of significance. In terms of initiation into cannabis, females are more likely to never use than males (54% compared to 46%), and less likely to initiate use before the age of 16 (8% compared to 9%), at age 16 or 17 (11% compared to 15%), or at 18 years of age or older (26% compared to 29%). The difference between males and females in terms of cannabis up-take are significantly different for each category.

The information in Table 2 can also be used to learn about the joint distribution

[7]The test statistic has a value of 45.04 for males and 106.75 for females compared to the critical value of 16.92 for a χ^2_9 at the 5% level of significance.

of educational attainment and age of initiation into cannabis use. For males, 43% of non-users have less than a high-school education compared to 47% of those who have used cannabis at some point in their life. Closer inspection reveals that the probability of having less that a high-school education amongst cannabis users is higher for those who initiate earlier. For example, 57% of males who initiate by the age of 15 fail to complete high-school compared with 48% of males who initiate between the ages of 16 and 17. At the other end of the education spectrum, while 16% of male non-users and 15% of male cannabis users have an undergraduate degree, the probability of a cannabis user graduating from university is much higher for those who initiate into use later, say after the age of 17 (16%) compared to those who initiate by the age of 15 (11%). This general pattern is common across all categories of educational attainment and across gender, providing some evidence that the effects of initiating into cannabis use differs by the age of initiation for both genders.

Due to the differences in decisions about starting cannabis use and leaving formal education across gender, the following analysis will be conducted separately for males and females.

4 Instrumental Variables Estimation

4.1 Econometric Set-up

We are interested in estimating the expected impact of initiation into cannabis use (c) on educational attainment (e). Following the previous literature, we specify our model for the production of education as:

$$e_i = \beta_1 + \delta_c c_i + x'_{1i}\beta_2 + x'_{2i}\beta_3 + \epsilon_i \tag{1}$$

where educational attainment of respondent i (measured by the age at which the respondent leaves school, an indicator for graduating from high-school, or an indicator for graduating from university), depends upon the observable variables, c_i, x_{1i} and

x_{2i}, and unobservable factors, ϵ_i. In addition to the effect of initiation into cannabis use (measured by δ_c), the production of educational attainment is assumed to depend upon a vector of exogenous individual demographic characteristics, x_1 (age, Australian born, state of residence and rural residence) and a vector of variables intended to capture permanent characteristics of the individual, such as the extent to which they discount the future or possess a "curious or adventurous" nature, x_2 (measured by an indicator for having ever smoked cigarettes daily by the time of survey and whether they have ever smoked cigarettes but not been a daily smoker by the time of survey, respectively). The random error term, ϵ_i, captures other unmeasured influences that affect the production of education, such as a taste for risk, or predisposition for non-conformist or deviant behavior (Jessor and Jessor, 1977).

As has been widely discussed in the economics literature, the challenge in ascertaining the causal impact of substance use on education outcomes is overcoming the possibility that substance use is endogenous. This endogeneity may result from academic failure leading to substance use (reverse causality) or from unobserved characteristics that lead teenagers to substance use also leading them to drop out of formal education (selection on unobservables or omitted variables). Failure to address the potential endogeneity of initiation into drug use will lead to biased and inconsistent estimates of its impact on educational attainment. As discussed in section 2, these issues have previously been tackled within the economics literature using an IV approach. In the first instance, we also follow this approach.

The ability of an IV estimator to consistently estimate the causal impact of cannabis use on educational attainment relies on the availability of variables that determine initiation into cannabis use but that can be validly excluded from the education production function. Previous research has sought to identify the effect of substance use, such as drinking, on education using state level policy variables related to substance use (Cook and Moore, 1993; Williams et al., 2003; Koch and McGeary, 2005). However, this identification strategy has come under criticism because (1) the state level policies are often only weakly correlated with substance use; and (2) state level policies may be correlated with unobserved state sentiment that influences both substance use and the economic outcomes of interest (Dee and Evans, 2003).[8] We propose to identify the effect of early initiation into cannabis use with the timing of having smoked a cigarette for the first time. To be more concrete, we measure early use of cannabis with an indicator for first use at or before the age of 15 and instrument this measure with an indicator for having first smoked a full cigarette by the age of 10, an indicator for having first smoked a cigarette at age 11, and an indicator for having first smoked a cigarette at age 12. This set of variables is intended to measure whether an individual has learned the technology to inhale smoke. As this technology is required for smoking cannabis, we expect that those who have previously tried smoking cigarettes are more likely to smoke cannabis compared to those who have not.[9] Moreover, having accounted for the effect of (permanent) individual characteristics such as a high discount rate or curiosity using indicators for ever been a daily smoker of cigarettes and having tried cigarettes but never smoked daily, the indicators for early cigarette use should not have a direct impact on educational attainment and can therefore be excluded from equation 1.[10]

If an IV procedure is to improve upon methodologies that treat drug use as exogenous it is vital that there exists a (sufficiently) strong correlation between

[8]Variables used to identify the impact of cannabis use on educational outcomes in the literature include: religiosity (Register et al., 2001), state level drug policies (cannabis decriminalization (Chatterji, 2006; Pacula at al. 2003; Register et al., 2001), criminal penalties for cannabis possession (Chatterji, 2006; Pacula et. al, 2005; Pacula at al. 2003)), the price of cannabis (Pacula et. al, 2005; Pacula et al., 2003), 8^{th} grade school characteristics (8^{th} grades school principal's perception of whether or not drugs are a moderate to serious problem, whether the school has a policy of expelling students if they are caught with drugs on school property (Chatterji, 2006) and siblings drug use (Pacula et. al, 2005).

[9]Over 80% of those who used cannabis in the past year reported smoking as the route of administration.

[10]We have, nonetheless, attempted to strengthen our case by only considering initiation into cigarette use up to three years before starting cannabis use.

the excluded instruments and the variables they are instrumenting. Failure of this condition is referred to as the problem of weak instruments. Unfortunately weak instruments have extremely deleterious effects on the sampling properties of IV estimators, inducing substantial bias in small samples and rendering invalid standard forms of inference such as t-tests and the construction of confidence intervals. Consequently, in what follows we pay considerable attention to the strength of our instruments to ensure that our results are not subject to this problem. We do this through examination of the first-stage regression and by conducting inference on the IV coefficients of interest using the Anderson-Rubin test, which is robust to the presence of weak instruments.

Estimation is carried out using the two-step efficient generalized method of moments (GMM). This method has the advantage of relaxing the assumption of identically and independently distributed error terms required of IV or two stage least squares estimation, and hence is robust to heteroscedasticity.

4.2 GMM Estimation Results

Table 3 contains the results from examining the relationship between initiation into cannabis by age 15 and educational attainment using OLS and GMM estimation. Each of the three columns in Table 3 report results based on a different measure of educational attainment. In the first column, educational attainment is measured by the age at which the respondent leaves formal education, in the second column it is measured by an indicator for completing high-school, and in the third column it is measured by an indicator for completing an undergraduate degree.[11] We present

[11]For space considerations, only the coefficient estimates on the measure of age of initiation into cannabis use and diagnostic statistics are reported in this table. All models control for the following characteristics: age, age squared, an indicator for ever being a daily smoker of cigarettes, an indicator for ever being a non-daily smoker of cigarettes, an indicator for being Australian born, an indicator for living in a rural area, and a set of indicators for the respondent's state of residence at time of survey. A complete set of results (excluding state fixed effects) based on the age at which the respondent left school can be found in Appendix Table A1.

separate estimates for the male and female sub-samples in the top and bottom panels, respectively. The reported t-statistics are based on heteroscedastic robust standard errors.

Before discussing coefficient estimates, we assess the statistical merits of the GMM estimates. We begin with the strength of the correlation between the measure of initiation into cannabis use and the instruments. For both males and females, the F-statistic for testing the joint significance of the instrument set in the first stage regression is well over the bench-mark value of 10 sometimes used to diagnose weak instruments (Staiger and Stock, 1997). This suggests that the issues associated with weak instruments are unlikely to be a problem in this analysis. Moreover, the first stage regression results show that each of the indicators for the age of first smoking a cigarette are individually statistically significant and correctly signed (see Appendix Table A1). The over-identification restrictions are also always supported for each gender, providing evidence that early cigarettes use (controlling for ever being a non-daily or daily cigarette smoker) are valid instruments for identifying the impact of early initiation into cannabis use.[12] Finally, the Hausman test finds significant differences between OLS and GMM estimates of the impact of initiation into drug use on the three measures of educational attainment for females, but not males. This suggests that while the OLS estimates are likely to suffer from endogeneity bias in the case of the females sub-sample, there is no significant evidence of this problem in the male sub-sample. On this basis, we prefer the GMM estimates in the case of females and OLS estimates in the case of males.

A comparison of the OLS and GMM coefficient estimates for females indicates that ignoring the endogeneity of early initiation into cannabis use results in an understatement of its negative impact on educational attainment. This suggests that females who are more likely to use cannabis by the age of 15 are also more likely to remain at school longer. Focusing now on the GMM parameter estimates,

[12] The Hansen J statistic is used to test the over-identifying restrictions. It is distributed as a χ^2-squared with 2 degrees of freedom.

the results in Table 3 suggest that early cannabis use by females reduces the expected age at leaving education by 1.8 years. Similarly, cannabis use by age 15 is expected to reduce the probability that a female completes high-school by 32 percentage points and the probability that she completes college by 21 percentage points. These are large effects given that on average, 55% of females in the sample graduate from high-school and 25% graduate from university. Moreover, the Anderson-Rubin test confirms that, accounting for the potential issue of weak instruments, these effects are statistically significant (although this is only true at the 7% level of significance in the case of completing university).

By comparison, the impact of early initiation into cannabis use has a smaller effect on the educational outcomes of males. First using cannabis by age 15 is expected to reduce the number of years of education completed by males by one third of a year, and reduce the probability of completing high-school and college by 10 and 4 percentage points, respectively. Nonetheless, these are still quite large effects given that 55% of the sample of males graduate from high-school and 26% graduate from university.

In terms of the impact of the control variables, we find that older individuals tend to complete fewer years of formal education, as do Australian born and rural respondents. We also find that individuals who have ever been a daily smoker of cigarettes tend to complete fewer years of education while those who have smoked cigarettes, but never on a daily basis, complete slightly more years of education compared to those who have never smoked cigarettes.

Overall, our findings from IV analysis are in broad agreement with those of the earlier studies, which conclude that early cannabis use reduces educational attainment. Moreover, averaging over the estimated effects of initiation into cannabis use for males and females, we find that cannabis use by the age of 15 is associated with approximately a 1.2 year reduction in years of education, compared to a range of 0.2 to 1 year for the earlier studies.

4.3 Sensitivity Analysis of the GMM Results

We examined the robustness of the GMM results to the potential for omitted variable bias and to the instrument set used to identify the impact of early initiation into cannabis use on educational attainment.[13]

So far, our analysis attempts to account for permanent characteristics of respondents, such as their discount rate or whether they have an adventurous nature by including measures of having ever been a daily smoker or ever been a non-daily cigarette smoker. It may also be argued that ones tendency to use alcohol also reflects these characteristics and therefore a measure of alcohol use should also be included in the education production function. Moreover, failing to do so may lead to omitted variable bias if alcohol use is correlated with both educational attainment and initiation into cannabis use. In order to examine this issue, we include an indicator for alcohol use in the past year in the education production function. This had very little impact on the point estimates and no impact on our qualitative findings, leading us to conclude that omitting current alcohol use has not biased our results.

The second robustness test we conduct examines the validity of our instrument set, which consists of a set of indicators for first smoking a full cigarette prior to the age of 13. This set of variables is intended to capture whether a person has acquired the technology to smoke cannabis. We argue that, while knowledge of such technology is likely to influence whether a person uses cannabis, it is unlikely to have any direct effect on their production of education. However, it may also be argued that early use of cigarettes signals deviant behavior. If this is the case, the exclusion restrictions may be invalidated. In order to investigate this hypothesis, we add to the list of instruments early alcohol use (indicators for first serve of alcohol consumed by the age of 10, at the age of 11, at the age of 12). While early alcohol use may signal deviant behavior in youth, it does not impart any technological knowledge useful for smoking cannabis. Therefore, we have no reason to believe that the set

[13] The estimation results are available on request.

of indicators for early alcohol use is a valid instrument set and statistical evidence supporting its exclusion from the education production function would cast doubt on our identification strategy based on early cigarette use. The over-identification test based on the instrument set consisting of early alcohol use and early cigarette use strongly rejects excluding this set of variables from the education production function. Since the instrument set based solely on early cigarette use comfortably passes the over-identification test, we infer that early cigarette use is not capturing youthful deviance that may also affect the production of education. On this basis we conclude that the set of indicators for early cigarette use serves as valid instruments.

Given the above findings, as a further robustness test we include both early and current alcohol use in the set of determinants of the production of education. While adding current alcohol consumption had very little impact on the estimated effect of age of initiation into cannabis on education, including indicators of early alcohol use increases the magnitude of the estimated effects. Nonetheless, it does not alter our basic finding, that early initiation into cannabis use has a significant and detrimental effect on educational attainment.

As a final check, we examined the robustness of our results to the estimation sample. In particular, we excluded observations on the 185 males and 193 females who started cannabis use before the age of 15, but who had left formal education prior to initiating into cannabis use. Removing these observations reduced the magnitude of the estimated effect of initiating cannabis use on leaving formal education, but does not otherwise alter our findings.

5 The Bivariate Duration Model

5.1 Econometric Set-up

The results presented in section 4 confirm the findings based on US data by Chatterji (2006), Pacula et at (2005) and Register, Williams and Grimes (2001), that early cannabis use reduces educational attainment. This raises the question of whether the adverse consequences of up-take at age 12 are the same as at age 16, and whether this effect persists at older ages. These questions relate to the issue of the timing of events, which is naturally handled within a duration model framework. Therefore, in this section, we model transitions into cannabis use and out of formal education using a bivariate mixed proportional hazard model in which the unobservable components of these transitions are potentially correlated. A major advantage of using the bivariate duration approach is that identification of the treatment effect does not rely on a conditional independence assumption and it is not necessary to have a valid instrument. Rather, identification comes from the timing of events, i.e. the order in which initiation into cannabis use and leaving formal education occurs. Given that economic theory does not suggest a natural instrument, this is a particularly useful feature of this approach.[14]

Beginning with initiation into cannabis use, we assume that the rate at which individuals start using cannabis is a function of the elapsed duration of time they are exposed to potential use, their observed characteristics, and their unobserved characteristics. Individuals are assumed to be at risk of initiating into cannabis from the age of 12, so the elapsed duration of time since exposure to potential use is age minus 12. The hazard of starting cannabis use at time t conditional on observed characteristics x and unobserved characteristics v is specified as:

$$\theta_c(t \mid x, v) = \lambda_c(t) \exp(x'\beta_c + v) \tag{2}$$

where $\lambda(t)$ represents individual duration dependence. The observed characteristics are the same as those contained in x_1 in the previous section, excluding age, and unobserved heterogeneity reflects differences in the susceptibility to the uptake of

[14] To study the impact of policy interventions on unemployment durations this so called 'timing-of-events' method is used in several studies. Abbring and Van den Berg (2003) give a formal proof of the identification of the treatment effect in a bivariate duration model. They show that in this framework, identification is achievable without the usual restrictions. An example of a study that applies this approach in drug research is Van Ours (2003) who studies whether or not cannabis is a stepping-stone for cocaine.

cannabis.[15] Duration dependence is modelled by using a flexible step function:

$$\lambda_c(t) = \exp(\Sigma_k \lambda_{ck} I_k(t)) \tag{3}$$

where k (= 1,..,15) indexes age intervals and $I_k(t)$ are time-varying dummy variables that are one in subsequent age intervals. Of the 15 age intervals, 14 are 1 year in length (age 12, 13, 14, .., 25) and the last interval is open: 25+ years. Because a constant term is estimated, λ_{c1} is normalized to zero. In estimation, we account for the fact that, at the time of survey, some individuals have not yet started to use cannabis but that they may do so later by allowing their duration until use to be right censored.

As a starting point, and to simplify notation, we assume that the hazard of leaving formal education at time t depends upon whether initiation into cannabis use has previously occurred, on the observable characteristics x, and unobserved characteristics u as follows:

$$\theta_s(t \mid x, I_c, u) = \lambda_s(t) \exp(x'\beta_s + \delta \ I_c + u) \tag{4}$$

where I_c is an dummy variable equal to 1 if initiation into cannabis occurred prior to or in the current period and 0 otherwise. Note that, if individuals start using cannabis and leave school at the same age it is assumed that initiation into cannabis use preceded school leaving. The parameter of interest is δ since it determines whether initiation into cannabis has a positive effect on school-leaving ($\delta > 0$), a negative effect on school-leaving ($\delta < 0$), or whether there is no relationship between the two ($\delta = 0$). In the empirical section we expand this model to allow for the impact of cannabis initiation to vary by the age at which initiation occurs. Duration dependence is again modelled using a flexible step function:

$$\lambda_s(t) = \exp(\Sigma_n \lambda_{sn} I_n(t)) \tag{5}$$

where n is a subscript for age-interval and $I_n(t)$ are time-varying dummy variables for one year age-intervals. In the first instance, we focus on school-leaving up to age

[15] Specifically, x includes an indicator for Australian born, an indicator for rural residence and a set of indicators for state of residence at the time of survey.

18, taking into account the censoring of durations for individuals who have not left school by age 18. We also consider school leaving up to age 23. Because a constant term is included in the model, λ_{s1} is normalized to 0. Note that the explanatory variables for the school-leaving rate are the same as those for the cannabis starting rate. Because we censor the duration of time in formal education (at age 18 in some specification and at age 23 in others) some individuals have completed durations of formal education and others have right-censored durations. This is accounted for in estimation.

The potential correlation between the unobserved components in the uptake of cannabis and school leaving hazards is taken into account by specifying the joint density function for the duration of non-use of cannabis t_c and the duration of time individuals spend at school t_s conditional on x as

$$f(t_c, t_s|x, I_c) = \int_u \int_v f_s(t_s|x, I_c, u) f_c(t_c|x, v) dG(u, v) \tag{6}$$

$G(u, v)$ is assumed to be a discrete distribution with 4 points of support (u^a, v^a), (u^a, v^b), (u^b, v^a), (u^b, v^b). The associated probabilities are denoted as follows: $\Pr(u = u^a, v = v^a) = p_1$, $\Pr(u = u^a, v = v^b) = p_2$, $\Pr(u = u^b, v = v^a) = p_3$, $\Pr(u = u^b, v = v^b) = p_4$. Here p_j $(j = 1, .., 4)$ is assumed to have a multinomial logit specification: $p_j = \frac{\exp(\alpha_j)}{\Sigma_j \exp(\alpha_j)}$ and the normalization is $\alpha_4 = 0$.

Correlation between the unobserved components of the cannabis starting rate and the school leaving rate indicates that there is an overlap in the susceptibility to cannabis use and the tendency to leave school. Perfect correlation would imply that susceptibilities overlap completely, in which case the distribution of the unobserved heterogeneity has just two points of support. The correlation could be either positive or negative. If positive, individuals who are susceptible to cannabis use are also susceptible to school leaving and individuals who are not susceptible to cannabis use are not susceptible to leaving school use either. If the correlation between the unobserved components is negative, individuals who are susceptible to cannabis use are not susceptible to leaving school and individuals who are not susceptible to cannabis use are susceptible to leaving school. If the correlation is not accounted for the estimated effect of cannabis use on school leaving will be biased. In case of positive correlation the effect will be overestimated, while in the case of negative correlation the effect will be underestimated.

5.2 Results from Estimation of the Bivariate Duration Model

Table 4 presents estimates of the bivariate duration model of the hazard of initiating into cannabis use and the hazard of leaving formal education in which initiation at different ages is constrained to have the same effect on the hazard of leaving education.[16] Table 5 reports models in which the impact of initiation is allowed to vary by the age at which it occurs. Separate models are estimated for males and females.

Because of the nature of our data, which consists of detailed information on school leaving by year of school up to the completion of high-school and information on the highest post-school qualification, Table 4 first reports estimates of specifications for leaving education up to age 18 (treating those who graduate from university with undergraduate or graduate degrees as having censored durations). Specifications in which we incorporate the information on post-school qualifications are reported under the heading of 'leaving formal education up to age 23'. The top panel of Table 4 reports estimates for models that permit correlation in the unobserved heterogeneity determining the hazard of starting cannabis and the hazard of leaving education. The bottom panel presents estimates of the impact of initiation into cannabis on the hazard of leaving education when the unobservables are assumed to be uncorrelated. The later estimates provides a benchmark for assessing the bias arising from failing to account for the potential correlation in unobserved characteristics affecting these decisions.

[16] To save space we do not report the parameter estimates for the state fixed effects included in both transition rates and the age dummies included in the cannabis starting rate. Appendix Table A2 provides a full set of results (excluding state fixed effects) for the bi-variate model focusing on the hazard of school leaving up to age 18.

We begin with a discussion of the results based on the specification examining school leaving up to age 18. As detailed in the previous section, we allow for the potential correlation in unobservables influencing cannabis up-take and the school-leaving using a flexible approach in which unobserved heterogeneity is assumed to follow a discrete distribution with four points of support. However, for both males and females, only two points of support could be empirically identified, implying that the unobserved components in the two transition rates are perfectly correlated. Inspection of the estimated points of support in each equation revealed that the unobserved heterogeneity in transition rates for cannabis up-take and school leaving are negatively correlated.[17] A likelihood ratio (LR) test reveals that this correlation in unobservables is only statistically significant for the female sub-sample, which is consistent with the results of the previous section.[18] This implies that females who are more likely to start using cannabis are also less likely to leave school while females that will never start using cannabis have a higher school-leaving rate. Ignoring the negative correlation between the two transition rates leads to an underestimate of the true causal effect of initiation into cannabis use on school-leaving for females (but not males). This is confirmed by the coefficient estimates in the bottom part of Table 4, which are based on models in which the correlation in the unobserved components of the two transitions are constrained to be equal to zero.

In terms of the distribution of unobserved heterogeneity, the estimates in Table 4 indicate that, conditional on the observed characteristics, 47% of females in the sample belong to the group with a positive cannabis starting rate and a low school-leaving rate while 53% belong to the group with a zero cannabis starting rate and a high school-leaving rate. The rate at which females who have used cannabis leave

[17]We identify one positive cannabis starting rate, and one that is equal to zero. Since the second mass-point in the cannabis starting rate is smaller than the first and the second masspoint in the school-leaving rate is larger than the first, the unobserved heterogeneity in transition rates for cannabis up-take and school leaving are negatively correlated.

[18]Note under the null hypothesis of no correlation between unobservables, the LR test statistic is distributed as a χ^2_1. The critical value for the test statistic at the 5% level of significance is 3.8.

formal education is 49% (exp(0.40)) greater than otherwise similar females who have not tried cannabis. Ignoring the correlation in the unobservables produces a downward biased estimate of the effect of initiation into cannabis use on the school-leaving rate of 22%.

For males, conditional on observed characteristics, 54% of the sample belong to the group who have a positive cannabis starting rate and 46% will never use cannabis. Given the lack of a significant correlation between the unobserved components of the transitions into cannabis use and out of formal education, it is not surprising that there is very little difference in the estimated effect of initiation into cannabis use on school leaving across models that do and do not account for correlation in unobservables. Table 4 shows that, conditional on unobserved characteristics, age and region of residence, males that start using cannabis have a lower school-leaving rate than males that did not do so. Starting to use cannabis increases males' school-leaving rate by around 27% (exp(0.24)).

In terms of the impact of observed characteristics, the parameter estimates in Table 4 show that cannabis starting rates are higher for males and females born in Australia. The peaks in the cannabis starting rate for males and females at ages 18, 20 and 25 (see Appendix Table A2) coincide with the peaks presented in Figure 1. School-leaving rates for males and females born in Australia and living in rural areas are higher than for their counterparts. The peak in the school-leaving rate at age 18 (see Appendix Table A2) coincides with the peak presented in Figure 3.

The second and fourth columns of Table 4 report parameter estimates for specifications in which school-leaving is modelled up to age 23. As shown in the table, the parameter estimates are not sensitive to whether we consider school-leaving up to 18 years or up to age 23. However, since the specification modelling school leaving up to age 23 is more informative about the impact of cannabis use later in youth on leaving post-school education, it forms the basis of the more detailed analysis contained in Table 5.

Table 5 presents the results from estimating an expanded bivariate duration model, in which the impact of initiation into cannabis use on leaving formal education is allowed to vary by the age at which initiation occurs. The point estimates suggest that there is a differential impact of starting cannabis use across ages for both males and females. This is confirmed by an LR test which rejects the null hypothesis of equal effects of each age of initiation for both males and females.

As shown in the first column of Table 5, initiating cannabis use up to age 17 has a significantly positive effect on the school-leaving rate of males, and from age 18 onwards it has no significant impact on the hazard of leaving formal education. We examined whether the impact of starting cannabis use was the same for males for each age from 12 to 17, and zero beyond the age of 17 using an LR test. As shown by the test statistic in the bottom part of Table 5, we are unable to reject this hypothesis at conventional levels of significance. Therefore, it appears that starting cannabis use before the age of 18 has a negative effect on educational outcomes for males, where the size of the effect on the school-leaving rate does not depend on the exact age of initiation.[19] Initiation into use after the age of 17 appears to have no effect on the rate at which males leave formal education.

By and large the effect of the age of initiation into cannabis use on the school-leaving rate of females is similar to those for males with two exceptions. First, initiation into cannabis use beyond the age of 17 appears to have an effect on the rate at which females leave school. Second, the magnitude of the estimated effect of initiation on school leaving seems to be larger for females than males. The restricted estimates and the corresponding LR test in the lower part of the second column of Table 5 confirm that initiation up to the age of 19 significantly increases the rate at which females leave formal education compared to those who have not yet started.[20]

[19]Note however that the effect in terms of the years of schooling is larger for those who start using cannabis at younger ages. This is because the effect of the increased school-leaving rate works over a longer calendar time period.

[20]Females that start using cannabis at age 18 have a significantly higher school-leaving rate than non-cannabis users, but the effect is significantly smaller than the effects related to starting at age 17 or age 19.

Therefore, the threshold age for females is 20 (compared to 18 for males). If females start using cannabis before the age of 20, it has a negative effect on their educational attainment, and the size of the effect on the rate at which they leave school does not (generally) depend on the exact age of initiation. Initiation into cannabis after the age of 19, however, appears to have no effect on the rate at which females leave formal education.

While the estimates presented in Table 4 and 5 address the potential for unobserved characteristics to jointly influence the decision to start using cannabis and the decision to leave formal education, they do not account for the potential influence of leaving formal education on cannabis up-take. We explored this issue by augmenting the model for initiation into cannabis as follows:

$$\theta_c(t|x, I_s, v) = \lambda_c(t) \exp(x'\beta_c + \gamma I_s + v) \quad (7)$$

where I_s is a dummy variable equal to 1 if the respondent left school prior to the current period and zero otherwise. The parameter γ determines whether leaving school impacts on the decision to initiate into cannabis use. The results from estimation indicate that leaving formal education has no (statistically significant) impact on the up-take of cannabis by males, but it does have a small positive impact on the up-take by females. Nonetheless, the estimated effect of initiation into cannabis on leaving formal education is not sensitive to accounting for reverse causality.[21]

In view of the non-linearity of the estimated model, we use the parameter estimates from the lower part of Table 5 to give a sense of the magnitude of the effect of starting cannabis use on the duration of formal education. On average sample members leave school at age 18. So, at age 13 boys and girls have about 5 years of education remaining and at age 15, they have an average of 3.5 years remaining. On

[21] For specifications that constrain the impact of initiation to be the same for all ages and consider leaving education up to age 23, the estimated impact of initiating cannabis use for males is 0.22 with a t-statistic of 5.8 when reverse causality is accounted for, compared to 0.22 with a t-statistic of 6.3 when it is ignored. For females, the estimated effect is 0.37 with a t-statistic of 10.8 compared to 0.36 with a t-statistic of 11.2 when reverse causality is ignored.

the basis of the parameter estimates from the bottom of Table 5, a boy that starts using cannabis at age 13 is expected to reduce his duration of time in formal education by 1.2 years. Starting cannabis use at age 15 reduces the expected number of years of formal education for males by 0.8 years. For girls initiating into cannabis use at age 13 and at age 15 reduces the expected number of years of education by 1.7 and 1.2 years, respectively. These estimates are remarkably consistent with those obtained from GMM estimation in the previous section.

6 Discussion

In this paper, we investigate the impact of initiation into cannabis use by youth on their educational attainment. Our contribution to the literature is threefold. First, in addition to traditional instrumental variable estimation, we use the dynamic framework of hazard rate analysis. Although each method has a different approach to addressing the potential endogeneity of cannabis use, the results are remarkably robust to the estimation strategy used. Second, we show that although self-selection into cannabis use does not appear to be an issue for males, it is for females. More specifically, we find that females who are inclined to start using cannabis are also more inclined to stay at school longer, while females who are unlikely to start using cannabis tend to leave school earlier. Failing to account for this correlation in unobserved components of the cannabis up-take and education leaving decisions leads to an underestimate of the effects of cannabis use on educational attainment. Our third contribution is to show that the magnitude of the effect of initiation into cannabis use depends on the age of onset. Using the bivariate duration framework, we found that those who initiate into cannabis use early suffer greater adverse effects in terms of educational attainment, whereas initiation at older ages – for males after age 17 and for females after age 19 – does not seem to have harmful effects.

Returning to the title of the paper, while we would prefer to offer worried parents some comfort, our findings tend to reinforce their concerns, at least with respect to educational attainment. However, it should also be noted that our results do not imply a mechanical relationship from cannabis use to poor educational achievement. Rather, our estimates show that early up-take of cannabis leads to a higher school-leaving rate, meaning that there is a probabilistic decrease in years of education completed. While our data are not rich enough to allow an investigation into why some youngster's initiation into cannabis use has a negative effect on their educational attainment while for others there is no effect, previous studies may provide some insights. Early initiation into cannabis use has been shown to lead to higher levels and longer duration of use by Pudney (2004) and Van Ours and Williams (2007), respectively. Further, Pacula et. al (2005) reports that it is frequent persistent use that leads to lower educational attainment. Therefore, it may be that those who start cannabis use at younger ages are more likely to be heavy persistent users, and it is this mode of use that has a deleterious effect on education.

Finally, we think that it is noteworthy that we find significant differences in the impact of early initiation into cannabis on educational attainment for males and females. Similar findings have been reported in the literature on the effect of alcohol use on educational attainment (Koch and McGeary, 2005). These findings are supported by medical evidence which suggests that women are more sensitive to alcohol related brain/cognitive damage than men (Nixon, 1994; Hommer et al., 2001). Given the similarities in compensating brain activity found in functional magnetic resonance imaging studies of heavy cannabis and alcohol users (Kananyama et al., 2004; Plefferbaum et al., 2001), we speculate that our results indicate a similar differential effect of cannabis on the cognitive abilities of males and females.

References

[1] Abbring, J.H, and Van den Berg, G.J. (2003) The Non-Parametric Identification of Treatment Effects in Duration Models, *Econometrica*, 71, 1491-1518.

[2] Bray, J. W., Zarkin, G. A., Ringwalt, C. and J. Qi (2000) The Relationship between Marijuana Initiation and Dropping Out of High School, *Health Economics*, 9(1), 9-18.

[3] Chatterji, P. (2006) Illicit Drug Use and Educational Attainment, *Health Economics*, 15(5), 489-511.

[4] Cook, P. and Moore, M.J. (1993) Drinking and Schooling, *Journal of Health Economics*, 12(4), 411-429.

[5] Dee, T.S. and Evans, W.N. (2003) Teen Drinking and Educational Attainment: Evidence from Two Sample Instrumental Variable Estimates, *Journal of Labor Economics*, 21(1), 178-209.

[6] Hommer, D.W., Momenan, R., Kaiser, E. and Rawlings, R.R. (2001) Evidence for a Gender-Related Effect of Alcoholism on Brain Volumes, *American Journal of Psychiatry*, 158(2), 198-204.

[7] Jessor, R and Jessor, S.L. (1977) *Problem behavior and psychological development: A longitudinal study of youth*. New York: Academic Press.

[8] Kananyama, G., Rogowska, J., Pope, H.G., Gruber, S. and Yurgelun-Todd, D.A. (2004) Spatial Working Memory in Heavy Cannabis Users: a Functional Magnetic Resonance Imaging Study, *Psychopharmacology*, 176, 239-247.

[9] Koch, S. F and McGeary, K.A. (2005) The Effect of Youth Alcohol Initiation on High-School Completion, *Economic Inquiry*, 43(4), 750-765.

[10] Macleod J., Oakes R., Copello A., Crome I., Egger M., Hickman M., Oppenkowski T., Stokes-Lampard H. and Smith G.D. (2004) Psychological and

social sequelae of cannabis and other illicit drug use by young people: a systematic review of longitudinal, general population studies, *The Lancet*, 363, 1579-1588.

[11] Nixon, S.J. (1994) Cognitive Deficits in Alcoholis Women, *Alcohol Health and Research World*, 18(3), 228-232.

[12] Pacula, R. L., Ellickson, P., Ringel, J. and Collins R. (2005) Revisiting the Relationship Between Marijuana and Educational Attainment, *mimeo*.

[13] Pacula, R.L., Ross, K.E. and Ringel, J. (2003) Does Marijuana Use Impair Human Capital Formation?, National Bureau of Economic Research, Inc, *NBER Working Papers* No. 9963.

[14] Plefferbaum, A., Rosenbloom, M., Deshmukh, A. and Sullivan, E.V. (2001) Sex Differences in the Effects of Alcohol on Brain Structure, *American Journal of Psychiatry*, 158(2), 188-197.

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

[16] Register, C. A., Williams, D. R. and P. W. Grimes (2001) Adolescent Drug Use and Educational Attainment, *Education Economics*, 9(1), 1-18.

[17] Staiger, D. and Stock, J.H. (1997) Instrumental Variables regression with weak instruments, *Econometrica*, 65, 557-86.

[18] Yamada, T., Kendix, M. and Yamada, T. (1996) The Impact of Alcohol Consumption and Marijuana Use on High School Graduation, *Health Economics*, 5(1), 77-92.

[19] Van den Berg, G.J. (2001) Duration models: specification, identification, and multiple durations, in: Heckman, J.J., and Leamer, E. (Eds.), *Handbook of Econometrics*, Volume V, Chapter 55, North-Holland, Amsterdam.

[20] Van Ours, J.C. (2003) Is cannabis a stepping-stone for cocaine?, *Journal of Health Economics*, 22, 539-554.

[21] Van Ours, J.C. and Williams, J. (2007) Cannabis Prices and Dynamics of Cannabis Use, *Journal of Health Economics*, forthcoming.

[22] Williams, J., Powell, L. and Wechsler, H. (2003) Does Alcohol Consumption Reduce Human Capital Accumulation? Evidence from the College Alcohol Study, *Applied Economics*, 35, 1227-1239.

Figure 1: Annual starting rate cannabis use (%)

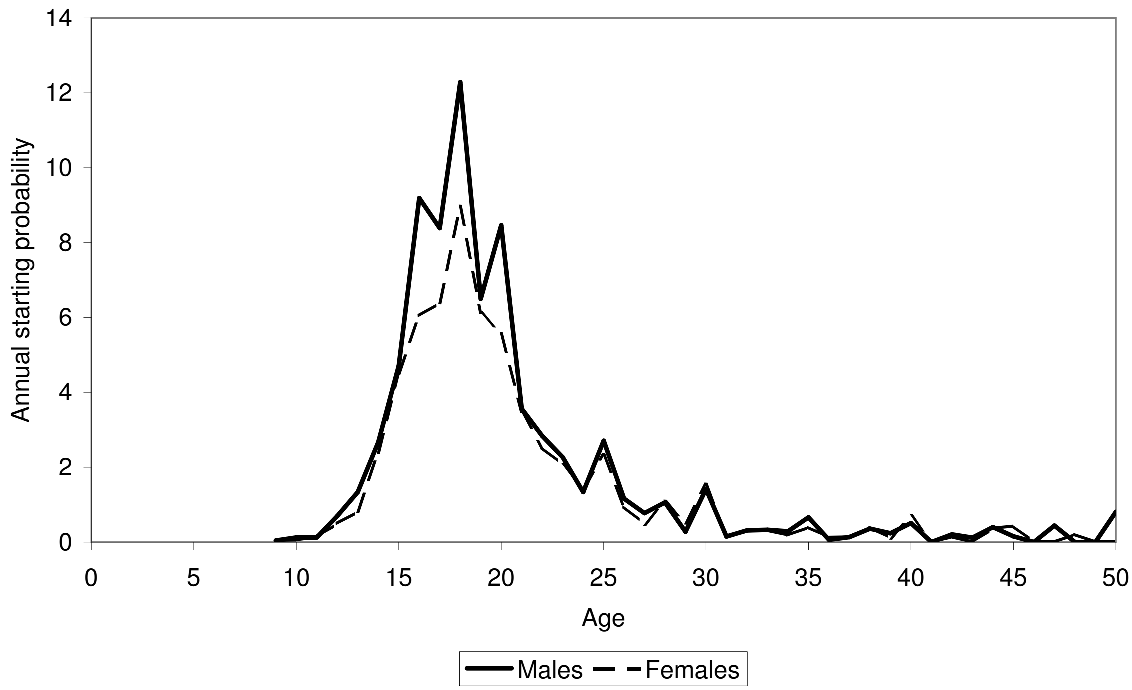

Figure 2: Cumulative starting probability cannabis use (%)

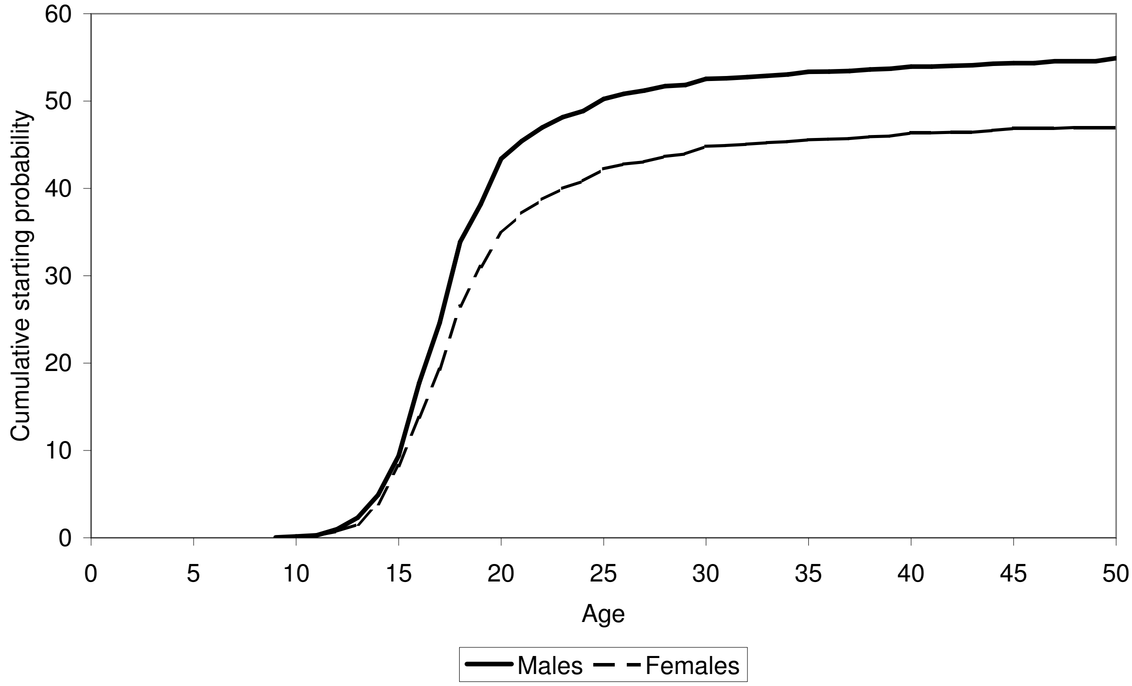

Figure 3: Annual school-leaving rate (%)

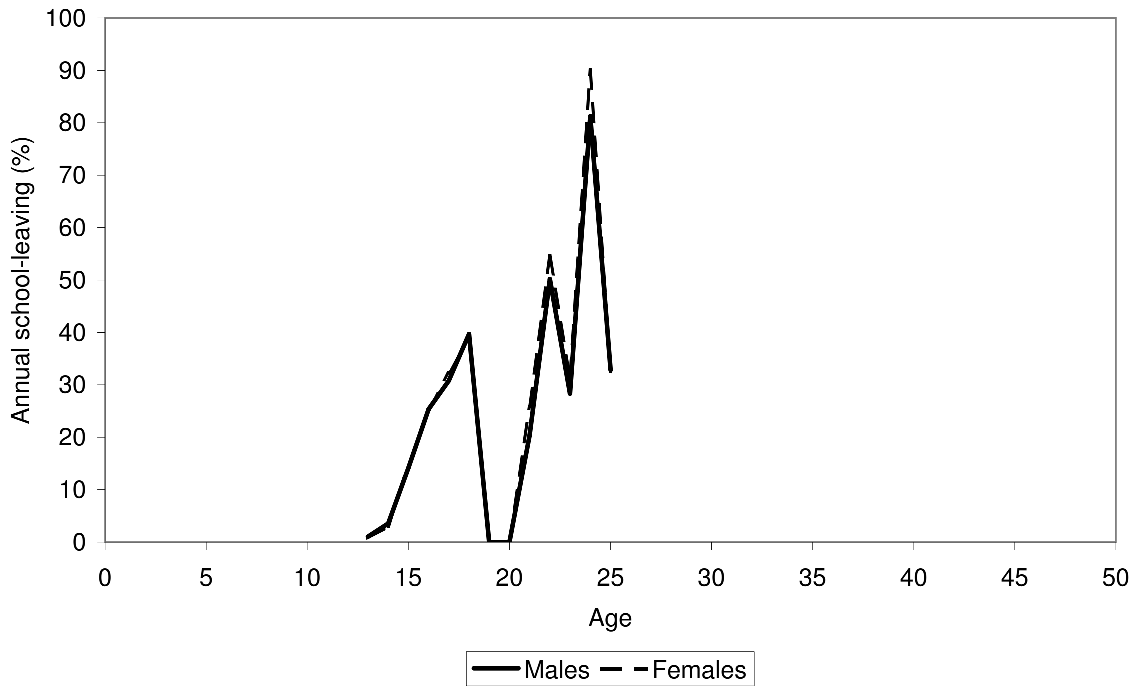

Figure 4: Cumulative probability of school-leaving (%)

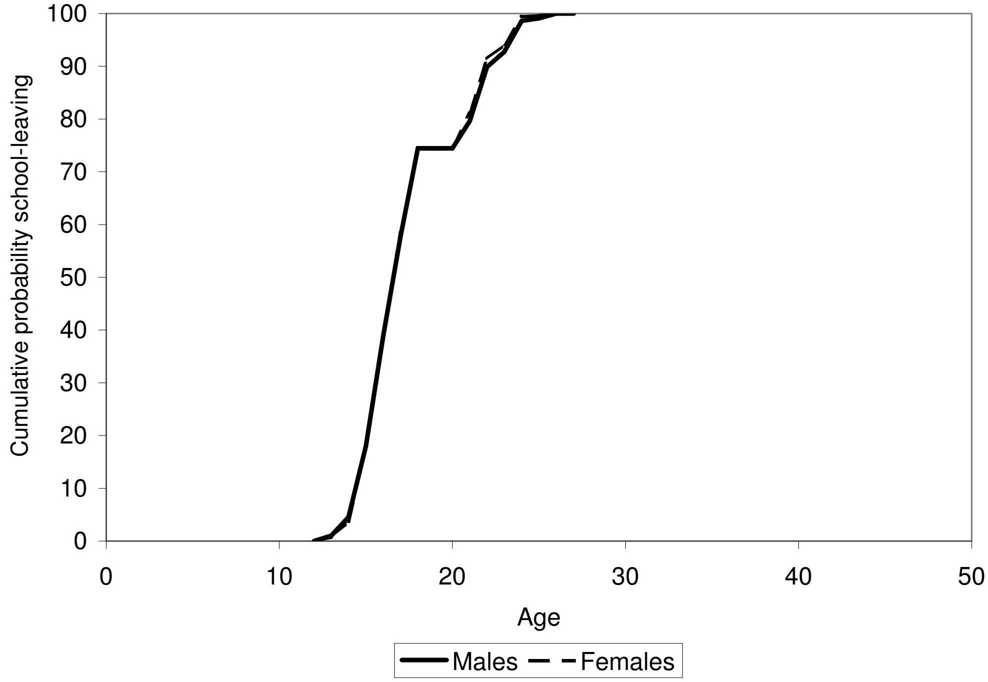

Table 1: **Summary statistics**

Variable	Mean	Min	Max
Cannabis use by age 15	0.09	0	1
Age of initiation (conditional on starting)	18.78	9	50
Age left school	17.94	13	27
Graduated from high school	0.55	0	1
Graduated from university	0.25	0	1
Male	0.42	0	1
Age	37.41	25	50
Born in Australia (Oz)	0.78	0	1
Rural	0.26	0	1
Ever smoked cigarettes daily	0.5	0	1
Ever smoked cigarettes but not daily	0.22	0	1
Smoked first cigarettes by age 10	0.05	0	1
Smoked first cigarettes by age 11	0.02	0	1
Smoked first cigarettes by age 12	0.07	0	1

Table 2: **Joint frequency distribution of initiation into cannabis use and educational attainment; males and females**

a. Males	Never use	Age<16	15<Age<18	Age>17	Total	Total (%)
< High-school	985	264	356	616	2221	45.2
High school	676	119	218	423	1436	29.3
Undergrad	364	49	119	225	757	15.4
Post-grad	255	29	55	159	498	10.1
Total	2280	461	748	1423	4912	100.0
Total (%)	46.4	9.4	15.2	19.0	100.0	
b. Females						
< High-school	1778	315	353	664	3110	45.2
High school	1063	145	238	575	2021	29.3
Undergrad	590	78	119	374	1161	16.9
Post-grad	302	29	51	207	589	8.6
Total	3733	567	761	1820	6881	100.0
Total (%)	54.3	8.2	11.1	26.4	100.0	

Table 3: **The effect of initiation into cannabis use by age 15 on age left school**

	Age left school	Complete high school	Complete university
a. Males			
Parameter estimates δ_c (absolute z-scores)			
OLS	-0.33 (2.5)*	-0.10 (4.2)*	-0.04 (2.0)*
GMM	-0.83 (0.9)	-0.36 (2.1)*	0.04 (0.3)
Test statistics (p-value)			
F-test on instruments	19.49 (0.00)	19.49 (0.00)	19.49 (0.00)
Over-identification	0.42 (0.42)	2.80 (0.25)	2.20 (0.33)
Hausman	0.25 (0.61)	2.39 (0.12)	0.41 (0.52)
Anderson-Rubin	2.51 (0.47)	7.54 (0.06)	2.29 (0.51)
b. Females			
Parameter estimates δ_c (absolute z-scores)			
OLS	-0.34 (3.0)*	-0.10 (4.5)*	-0.02 (1.3)
GMM	-1.85 (3.0)*	-0.32 (2.8)*	-0.21 (2.2)*
Test statistics (p-value)			
F-test on instruments	35.12 (0.00)	35.12 (0.00)	35.12 (0.00)
Over-identification	2.28 (0.32)	1.49 (0.47)	2.02 (0.36)
Hausman	5.36 (0.02)	3.82 (0.05)	3.38 (0.07)
Anderson-Rubin	11.70 (0.01)	9.33 (0.03)	6.97 (0.07)

Note: 4912 males and 6881 females; all models control for the following characteristics: gender, age, age squared, an indicator for ever being a daily smoker of cigarettes, an indicator for ever being a non-daily smoker of cigarettes, an indicator for being Australian born, an indicator for living in a rural area, and a set of indicators for the respondent's state of residence at time of survey; a * indicates that the coefficient is different from zero at a 5% level of significance.

Table 4: **Parameter estimates bivariate duration models – starting rates cannabis use and school-leaving rates; school-leaving up to 18 years and up to 23 years**

	Males		Females	
a. Correlation	**≤ 18 years**	**≤ 23 years**	**≤ 18 years**	**≤ 23 years**
Cannabis use				
Oz	0.22 (4.2)*	0.22 (4.2)*	0.37 (6.9)*	0.37 (6.9)*
Rural	-0.03 (0.7)	-0.04 (0.7)	-0.06 (1.4)	-0.06 (1.4)
Masspoint 2	$-\infty$	$-\infty$	$-\infty$	$-\infty$
School-leaving				
Oz	0.35 (10.1)*	0.35 (11.4)*	0.25 (8.1)*	0.27 (10.3)*
Rural	0.54 (17.9)*	0.47 (17.6)*	0.30 (12.0)*	0.31 (13.5)*
Masspoint 2-1	0.03 (0.9)	0.05 (1.5)	0.32 (11.1)*	0.32 (11.3)*
Effect cannabis use (δ)	0.24 (6.4)*	0.22 (6.3)*	0.40 (11.1)*	0.36 (11.2)*
Unobs. heterogeneity (α)	0.18 (5.9)*	0.18 (5.9)*	-0.12 (4.7)*	-0.12 (4.6)*
Probability (p_1)	0.54	0.54	0.47	0.47
-Loglikelihood	17,392.0	18,692.5	23,373.8	25,041.3
b. No correlation				
Effect cannabis use (δ)	0.23 (7.2)*	0.19 (7.0)*	0.20 (6.7)*	0.15 (5.9)*
-Loglikelihood	17,392.3	18,693.2	23,418.8	25,088.6
LR-test no correlation	0.6	1.4	90.0*	94.6*

The datasets contain 4912 males and 6881 females; Table 4a concerns parameter estimates with correlated unobserved heterogeneity; Table 4b has the same set-up except for the correlation between the unobserved heterogeneity terms which is ignored; all estimates include territories fixed effects (7) both in the cannabis use starting rate and in the school-leaving rate; note that the starting rates for cannabis use contains 14 age dummies (annually 13-25 and 25+ years); the rates for leaving school up to 18 years contain 5 age dummies (annually 14-18); the rates for leaving school up to 23 years contain 8 age dummies (annually 14-18 and 21-23); absolute t-statistics in parentheses; a * indicates that the coefficient is different from zero at a 5% level of significance.

Table 5: **Parameter estimates bivariate duration models – starting rates cannabis use and school-leaving rates; school-leaving up to 18 years – sensitivity analysis effect cannabis use**

	Males	Females
δ_{12-13}	0.44 (6.2)*	0.43 (5.7)*
δ_{14}	0.37 (4.6)*	0.45 (7.0)*
δ_{15}	0.23 (4.1)*	0.50 (10.1)*
δ_{16}	0.23 (4.8)*	0.32 (6.5)*
δ_{17}	0.19 (2.7)*	0.42 (7.4)*
δ_{18}	0.05 (0.7)	0.15 (2.4)*
δ_{19}	0.19 (1.5)	0.47 (5.8)*
δ_{20}	-0.09 (0.5)	0.14 (1.5)
δ_{21}	0.05 (0.3)	0.01 (0.1)
δ_{22-23}	0.15 (0.8)	-0.05 (0.3)
-Loglikelihood	18,683.40	25,026.70
LR-test equal δ's	18.6*	29.2*
Restrictions		
δ_{12-17}	0.26 (7.8)*	-
δ_{18-23}	0 (-)	-
$\delta_{12-17,19}$	-	0.42 (13.1)*
δ_{18}		0.15 (2.3)*
δ_{20-23}	-	0 (-)
-Loglikelihood	18,687.7	25,029.9
LR-test restrictions	8.6	6.4

The datasets contain 4912 males and 6881 females; all estimates include the same explanatory variables, fixed effects and age dummies as in Table 4. The first two and last two age-of-onset categories consist of two years because of the limited number of observations. Note that for females the LR-test of imposing $\delta_{12-17,19}$=δ_{18} equals 11.8, so we cannot reject the hypothesis that the effect of cannabis uptake at age 18 is different; absolute t-statistics in parentheses; a * indicates that the coefficient is different from zero at a 5% level of significance.

Table A1: GMM Parameter estimates of the effect of starting cannabis by age 15 on age left school

	Males	Females
Cannabis use by the age of 15		
Ever smoked cigarettes daily	0.11 (13.3)*	0.10 (16.2)*
Ever smoked cigarettes but not daily	0.02 (2.6)*	0.01 (2.1)*
Age	-0.02 (2.4)*	-0.01 (2.2)*
Age squared	0.00 (1.3)	0.00 (1.1)
Oz	0.02 (2.2)*	0.00 (0.10)
Rural	-0.01 (1.3)	0.00 (0.4)
Smoked first cigarettes by age 10	0.10 (4.6)*	0.18 (6.0)*
Smoked first cigarettes by age 11	0.11 (3.2)*	0.13 (3.9)*
Smoked first cigarettes by age 12	0.12 (5.7)*	0.17 (7.6)*
Constant	0.45 (3.6)*	0.33 (3.4)*
$\overline{R}^2$	0.10	0.11
Age left school		
Cannabis use by the age of 15	-0.83 (0.9)	-1.85 (3.0)*
Ever smoked cigarettes daily	-1.06 (6.6)*	-0.84 (7.7)*
Ever smoked cigarettes but not daily	0.25 (2.0)*	0.31 (3.1)*
Age	-0.11 (1.8)	-0.06 (1.3)
Age squared	0.00 (1.6)	0.00 (0.1)
Oz	-0.81 (7.6)*	-0.54 (6.3)*
Rural	-0.93 (10.6)*	-0.61 (8.4)*
Constant	22.08 (19.1)*	21.65 (23.2)*
$\overline{R}^2$	0.17	0.12

The datasets contain 4912 males and 6881 females; all estimates include state fixed effects; absolute z-scores in parentheses; a * indicates that the coefficient is different from zero at a 5% level of significance.

Table A2: Parameter estimates bivariate duration model – school-leaving up to 18 years

	Males	**Females**
Cannabis use		
Oz	0.22 (4.2)*	0.37 (6.9)*
Rural	-0.03 (0.7)	-0.06 (1.4)
Age 13	0.33 (1.7)	0.11 (0.6)
Age 14	1.06 (6.3)*	1.23 (7.7)*
Age 15	1.70 (10.8)*	1.95 (13.0)*
Age 16	2.53 (16.8)*	2.35 (16.0)*
Age 17	2.55 (16.8)*	2.53 (17.2)*
Age 18	3.22 (21.7)*	3.12 (21.6)*
Age 19	2.69 (17.2)*	2.90 (19.6)*
Age 20	3.28 (21.4)*	3.02 (20.3)*
Age 21	2.53 (14.8)*	2.71 (17.2)*
Age 22	2.46 (13.8)*	2.54 (15.4)*
Age 23	2.39 (12.8)*	2.48 (14.4)*
Age 24	1.96 (9.0)*	2.23 (11.9)*
Age 25	2.95 (16.4)*	2.98 (17.3)*
Age 25+	2.12 (12.3)*	2.25 (13.6)*
Masspoint 1	-3.98 (24.0)*	-4.35 (27.4)*
Masspoint 2	$-\infty$	$-\infty$
School-leaving		
Oz	0.35 (10.1)*	0.25 (8.1)*
Rural	0.54 (17.9)*	0.30 (12.0)*
Age 14	1.98 (12.3)*	2.11 (13.5)*
Age 15	3.52 (24.0)*	3.86 (26.9)*
Age 16	4.27 (29.3)*	4.53 (31.6)*
Age 17	4.65 (32.1)*	4.96 (34.8)*
Age 18	4.96 (34.1)*	5.15 (36.0)*
Masspoint 1	-5.81 (38.0)*	-5.94 (39.7)*
Masspoint 21	0.03 (0.9)	0.32 (11.1)*
Effect cannabis use (δ)	0.24 (6.4)*	0.40 (11.1)*
Unobs. heterogeneity (α)	0.18 (5.9)*	-0.12 (4.7)*
Probability (p_1)	0.54	0.47
-Loglikelihood	17,392.0	23,373.8

The datasets contain 4912 males and 6881 females; all estimates include state fixed effects; absolute t-statistics in parentheses; a * indicates that the coefficient is different from zero at a 5% level of significance.