CANNABIS USE AND MENTAL HEALTH PROBLEMS

JAN C. VAN OURS[a,b,c]* AND JENNY WILLIAMS[b]

[a] *Department of Economics, CentER, Tilburg University, The Netherlands*
[b] *Department of Economics, University of Melbourne, Parkville, Australia*
[c] *CEPR, London, UK*

SUMMARY

This paper investigates whether cannabis use leads to worse mental health. To do so, we account for common unobserved factors affecting mental health and cannabis consumption by modeling mental health jointly with the dynamics of cannabis use. Our main finding is that using cannabis increases the likelihood of mental health problems, with current use having a larger effect than past use. The estimates suggest a dose–response relationship between the frequency of recent cannabis use and the probability of currently experiencing a mental health problem. Copyright © 2010 John Wiley & Sons, Ltd.

1. INTRODUCTION

Cannabis is the most widely used illicit drug. Over the last 30 years, the age at which it is first used has fallen and lifetime prevalence has risen in most developed countries (Hall, 2006). Cannabis's popularity is derived from the mild euphoria associated with its consumption and from the generally held belief that its health consequences are rather benign. However, there is growing evidence of an association between mental health problems and cannabis use. What remains unclear is whether the proper interpretation of this evidence is that cannabis use causes mental health problems. The existence of unobserved personal characteristics or circumstances that causes both mental illness and cannabis use is a plausible alternative explanation. The purpose of this paper is to investigate the nature of the relationship between cannabis consumption and mental health and, in so doing, determine the extent to which cannabis use leads to worse mental health.

Establishing whether cannabis use is a cause of mental illness is of particular interest from a policy perspective. Uptake of cannabis typically occurs during the mid to late teens while individuals are still attending school. For example, 42% of 12th graders in the USA and 32% of 12th graders in Australia have used cannabis in their lifetime (Johnston *et al*., 2006; White and Hayman, 2006). If cannabis use is a cause of mental illness, then educating adolescents about this risk may deter its uptake and thereby reduce population levels of mental illness. While a reduction in the prevalence of mental illness is desirable in itself, it is also likely to lead to significant economic benefits. The World Health Organization estimates the economic cost of mental illness to be between 3% and 4% of GNP per year for developed countries, with around half of the cost attributed to lost productivity (World Health Organization, 2003). Knowledge of the mental health consequences of cannabis consumption is also useful for informing the debate over its legal status

---

* Correspondence to: Jan C. van Ours, Department of Economics, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: vanours@uvt.nl

because it would allow a more accurate accounting of the costs and benefits of maintaining its status as a criminal offense. Given the recent moves to legalize cannabis in England and Portugal, this issue is clearly of ongoing policy interest.

There exists a substantial literature in economics documenting the consequences of cannabis use in terms of educational attainment, physical health and labor market success (for overviews see Van Ours and Williams, 2009; Williams and Skeels, 2006). Previous research on the relationship between mental health and illicit substance use, however, comes almost entirely from epidemiology.[1] The earliest attempt to identify the causal impact of cannabis use on mental illness is by Andreasson *et al.* (1987), who study a cohort of more than 50,000 18- to 20-year-old Swedish conscripts. The authors find that the post-conscription risk of developing schizophrenia is increasing in the number of times cannabis is used prior to conscription. Giving a causal interpretation to this finding is complicated, however, by the fact that while the prevalence of cannabis use has increased over the last 30 years in most countries, the prevalence of schizophrenia has not (Hall, 2006; Kalant, 2004). Nonetheless, Andreasson *et al.*'s (1987) research has prompted a raft of epidemiological studies on the topic. These subsequent studies tend to consider more general measures of mental health, including depression and anxiety. The results from this research are mixed, with some papers reporting a positive association between cannabis use and mental health problems (Fergusson *et al.*, 2002, 2005; Patton *et al.*, 2002; Rey *et al.*, 2002; Boys *et al.*, 2003) and others reporting no association (McGee *et al.*, 2000; Fergusson *et al.*, 1997). In a meta-analysis, Degenhardt *et al.* (2003) found a modest but significant association between heavy use of cannabis and later depression. In their overview study, Arseneault *et al.* (2004) conclude that rates of cannabis use are approximately twice as high among people with schizophrenia than among the general population.

In examining the relationship between mental health and cannabis use, the literature cited above has attempted to identify the causal effect of cannabis use by controlling for observed factors that may be a source of confounding.[2] However, as noted by Pudney (2010), the potential for unobserved common confounding factors makes inference regarding the causal impact of cannabis use difficult. The purpose of this paper is to address this issue. To do so, we use a discrete-factor approach. Our methodology marries Heckman and Singer's (1984) use of discrete factors in addressing unobserved heterogeneity in duration models with their use by Mroz (1999) to account for endogenous variables in regression models. More specifically, we estimate a trivariate system of equations consisting of hazard functions for the decision to start using cannabis and the decision to quit, and a Tobit model for the production of mental health. By allowing the distribution of discrete factors determining cannabis use dynamics and continuous mental health to be correlated, we account for common unobserved factors and hence obtain reliable estimates of the impact of cannabis use on mental health.

Our main finding is that frequent use of cannabis increases the likelihood of mental health problems. Infrequent and past cannabis use also increases the likelihood of mental health problems but the effects are substantially smaller. To give a sense of the magnitude of the effects, our estimates suggest that 2.4% of males who use cannabis weekly or more often will experience severe mental health problems compared with 1.5% of males who use monthly, 1.4% of males who are past users and 0.9% of males who have never used cannabis.

---

[1] For interesting reviews see Hall (2006), Kalant (2004), and Macleod *et al.* (2004).
[2] Fergusson *et al.* (2002, 2005) are exceptions in that they also control for unobserved heterogeneity using fixed-effects models. The later paper addresses the issue of causality using models estimated with LISREL.

The rest of this paper is laid out as follows. Section 2 describes the data used in this study and discusses their strengths and weaknesses. Section 3 presents the econometric methodology and results from estimation. Section 4 reports on an examination of the robustness of the results by way of an extensive sensitivity analysis. Section 5 summarizes our findings.

2. DATA

2.1. Australian National Drug Strategy Household Survey

This research draws on information collected in the 2004 Australian National Drug Strategy Household Survey (NDSHS). The NDSHS is managed by the Australian Institute of Health and Welfare on behalf of the Commonwealth Department of Health and Ageing. It is designed to provide data on awareness, attitudes and behavior relating to licit and illicit drug use by the non-institutionalized civilian population in Australia. The sampling framework is a multistage stratified sample design, where stratification is based on geographic region. In each sampled household, the respondent is the person with the next birthday who is at least 12 years of age. Self-completion questionnaires and computer-assisted telephone interviewing methodologies were used to survey respondents, with the bulk of data (82%) collected by self-completion questionnaires. Of the households contacted for the Drop and Collect Survey who fell within the scope of the study, 68% agreed to participate and accepted a questionnaire. Seventy-nine percent of these households returned the questionnaire. However, only 66% of the returned surveys were deemed usable, with the balance of questionnaires returned blank (8%) or with missing essential information or otherwise unreliable (5%). An analysis of the profile of non-response based on those who returned blank Drop and Collect Questionnaires indicates that this form of non-response was most prevalent among the over-60 age group, especially females over 60 (Australian Institute of Health and Welfare and the Australian Government Department of Health and Ageing, 2005).

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 an 'objective' measure of mental health in the form of the Kessler-10 (K10) scale of psychological distress, make these data useful for examining the impact of cannabis use on mental health outcomes.

2.2. Cannabis Use, Mental Health and Data Issues

Several measures related to cannabis consumption are used in our analysis. In modeling cannabis use dynamics, the outcomes of interest are the age at which cannabis was first used and the duration of use. The age of first use is constructed from responses to the question: 'About what age were you when you first used marijuana/cannabis?' This question was asked of all those who reported ever using cannabis. While we do not have information on the age at which respondents last used cannabis and hence duration of use, we do know whether or not they have used in the year prior to survey. Uncertainty surrounding the duration of cannabis use is addressed using econometric techniques, which are described in Section 3.

In modeling the production function for mental health, current and past use of cannabis are the focus. These measures are directly related to the outcomes in the dynamics of cannabis use. We

define current cannabis users as those who have initiated into cannabis and have used at least once in the 12 months prior to survey. Past users are defined as those who have initiated use but have not consumed cannabis in the 12 months before being surveyed. For current users, we also investigate whether the mental health impacts differ by the frequency with which cannabis is used. To do this, current cannabis users are categorized as using once or twice a year, every few months, about once a month, once a week or more, or every day. These mutually exclusive categories form the set of potential responses to the question: 'In the last 12 months, how often did you use marijuana/cannabis?'

The measure of mental health we use is the K10 scale of psychological distress. The K10 was developed as a screening tool for non-specific psychological distress for the US National Health Interview Survey (Kessler *et al.*, 2002, 2003). The K10 is widely used, both as a measure of mental health status in general population surveys and as an outcome measure in primary care settings (Pirkis *et al.*, 2005). It has been shown to be significantly correlated with other instruments, including the General Health Questionnaire, the Short Form 12, the Comprehensive International Diagnostic Interview–Short Form, and the World Health Organization Disability Assessment Schedule (Kessler *et al.*, 2003; Andrews and Slade, 2001). The K10 is based on self-reported responses to 10 items which ask respondents about symptoms of depression and anxiety experienced in the previous 4 weeks. It uses a five-level response scale that ranges from none of the time (1) to all of the time (5). The specific items asked are as follows: 'In the last 4 weeks, about how often did you feel . . .': 1. Tired out for no good reason? 2. Nervous? 3. So nervous that nothing could calm you down? 4. Hopeless? 5. Restless or fidgety? 6. So restless you could not sit still? 7. Depressed? 8. That everything was an effort? 9. So sad that nothing could cheer you up? 10. Worthless? The sum of scored responses to the 10 questions is used to generate a single 'score' of psychological distress ranging from 10 to 50. A score under 20 indicates that the respondent is likely to be well, a score of 20–24 indicates a mild mental disorder, a score of 25–29 indicates a moderate disorder and a score of 30 or greater indicates a severe mental disorder.[3]

As with previous research studying transitions in substance use with cross-sectional data, this study is subject to potential measurement error problems associated with recall error. As discussed below, we find some evidence of recall error in the reported age of initiation into cannabis use for those who report initiating after the age of 25. To the extent that respondents make errors in the age they report first using cannabis, our parameter estimates are likely to be biased towards zero. Since initiation into use after the age of 25 is fairly rare, we do not anticipate large effects from this source of measurement error. Nonetheless, we investigate the impact of this issue on our results in a sensitivity analysis contained in Section 4.

A further potential measurement issue relates to the use of the K10 score as the measure of mental health. The K10 score items relate to symptoms experienced in the 4 weeks prior to survey. If symptoms of mental illness are not experienced in the relevant 4-week window, then they may go undetected by this measure.

A more serious shortcoming of the data used in our study is that, while retrospective information is collected about the age when cannabis was first used, no such information is available for the age at which symptoms of mental illness were first experienced or the age at which mental illness

---

[3] These cut-offs are recommended for Australian General Practitioners (see http://www.gpcare.org/outcome%20measures/outcomemeasures.html).

was first diagnosed.[4] Therefore, we are unable to account for the potential for mental illness preceding, or causing, cannabis use. Several studies from epidemiology have, however, sought to determine whether there is a causal pathway running from mental illness to cannabis use. For example, Fergusson *et al*. (2005) investigate the relationship between cannabis use and psychotic episodes measured at ages 18, 21 and 25. They found no evidence that psychotic episodes lead to cannabis use and some evidence that increasing psychotic symptoms were associated with a decline in the use of cannabis. These findings are in line with Van Os *et al*. (2002) and Henquet *et al*. (2005), who find no evidence that early psychotic symptoms predict an increased risk of cannabis use, and Patton *et al*. (2002), who find that anxiety in teenagerhood does not predict later cannabis use. McGee *et al*. (2000), however, finds that although cannabis use at age 18 predicts mental disorders at age 21, mental disorders at age 15 predicts a small but significantly elevated risk of cannabis use at age 18. So, while we cannot rule out the possibility of reverse causality, the evidence suggests that it is unlikely to have a large effect on our estimates.

2.3. Descriptive Statistics

Our sample is composed of 4771 males and 6719 females aged 26–50 years old, for whom we have complete data on mental health, cannabis use and the other control variables. Summary statistics for the outcomes of interest and other explanatory variables are reported in Table I. Table I shows that 58% of males and 49% of females in the sample have used cannabis in their lifetime. Twenty percent of males and 11% of females have used cannabis in the past year. Therefore, according to our definitions, 38% of males and females are past users of cannabis, while 20% of males and 11% of females are current users of cannabis. Among those who have ever used cannabis, the average age of initiation is 18.4 years for males and 18.7 years for females, with 12% of males and 9% of females initiating before the age of 16. The frequency of past year use is also reported in Table I. For comparability with other variables, these rates are reported as percentages of the full sample. As shown in Table I, 5.2% of males and 3.7% of females in our sample use cannabis once or twice a year, 2.9% of males and 1.5% of females use every few months, 2.2% of males and 1.3% of females use every month, 5.3% of males and 2.4% of females use once a week, while 4.6% of males and 1.7% of females use cannabis every day. In terms of mental health, the average K10 score for males is 14.8 and for females it is 15.5. On the basis of their K10 score, 14% of males and 17% of females have a mental health disorder.[5]

In terms of demographic characteristics, the average age of the sample is about 38, and close to 80% of the sample were born in Australia, with 2% identifying themselves as Aboriginal or Torres Strait Islanders. More than two-thirds of individuals in the sample are currently married, and a further 10% of males and 14% of females are divorced. In terms of education, 16% of males

---

[4] As far as we have been able to determine, this weakness is not overcome in any other data. For example, the National Longitudinal Survey of Youth 97 cohort (NLSY97) uses a five-item short version of the Mental Health Inventory (MHI) to first assess mental health (in the past month) when respondents are aged 15–20, and the Christchurch Health and Development Study (CHDS) first asks about anxiety and depression at age 15. Thus neither study identifies the age at which mental health problems first occur.

[5] The Australian 2007 National Survey of Mental Health and Wellbeing collected information for the K10 as well as diagnosing mental disorders based on the International classification of Disease–10th Revision, Classification of Mental Health and Behavioral Disorders. It finds a high correlation between the two instruments, with 80% of those with a K10 of 30 or more being diagnosed with having a mental disorder in the previous 12 months, compared with only 11% with a score of less than 15 being diagnosed as having a mental disorder in the past 12 months (Australian Bureau of Statistics, 2008).

Table I. Means of variables

Variable	Males	Females
Ever used cannabis	0.58	0.49
Age first used cannabis (conditional on ever using)	18.4	18.7
Cannabis use before age 16	0.12	0.09
Cannabis use in past 12 months	0.20	0.11
Intensity of cannabis use in past 12 months (%)		
Once or twice a year	5.2	3.7
Every few months	2.9	1.5
About once a month	2.2	1.3
Once a week or more	5.3	2.4
Every day	4.6	1.7
K10 score	14.8	15.5
K10 indicator for mental health problem	0.14	0.17
Age	38.4	37.9
Born in Australia	0.76	0.79
Aboriginal or Torres Strait Islander	0.02	0.02
Married	0.68	0.70
Divorced	0.10	0.14
Highest qualification is year 10 certificate	0.16	0.24
Victoria	0.22	0.23
Queensland	0.18	0.19
South Australia	0.08	0.08
Western Australia	0.11	0.10
Tasmania	0.04	0.04
Australian Capital Territory	0.05	0.04
Northern Territory	0.05	0.05
Lives in a capital city	0.65	0.64
N	4771	6719

and 24% of females report their highest level of educational attainment as a 10th-grade education or less.

Figure 1 provides a graphical illustration of the probability of starting cannabis use at each age, conditional on not using up to that age. The figure shows that initiation into cannabis use begins at an early age. The first peak in the probability of uptake is at age 16, when 10% of males and 8% of females who had not previously used cannabis initiate use. The mean peak is at age 18, with 14% of males and 11% of females initiating use, but there is also a peak of 10% for males at age 20 and subsequent peaks are at age 25 and 30. The peaks at ages 25 and 30 in the age-specific starting probabilities may point to bundling in the recollection of the starting age although, as Figure 1 shows, initiation into cannabis use rarely occurs beyond age 25. In terms of the cumulative distribution, by the age of 16, 12% of males and 9% of females in the sample have started cannabis use. This increases to 54% of males and 45% of females at age 25. By the age of 50, 58% of males and 49% of females have used cannabis at some point in their life.

Figure 2 shows the distribution of K10 scores for males and females. By construction, the K10 score has a lower bound of 10 and an upper bound of 50, with higher scores indicating greater levels of psychological distress. The K10 score has a right-skewed distribution, with an average value around 15 for both males and females. A large proportion of observations (17% of males and 11% of females) occur at the lower bound value of 10. Figure 2 also shows that the sample proportion with each score falls as the score increases. As mentioned above, scores below
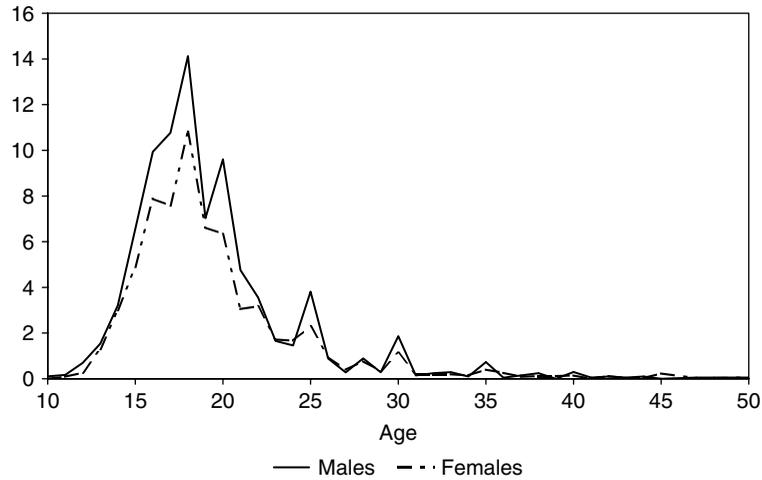
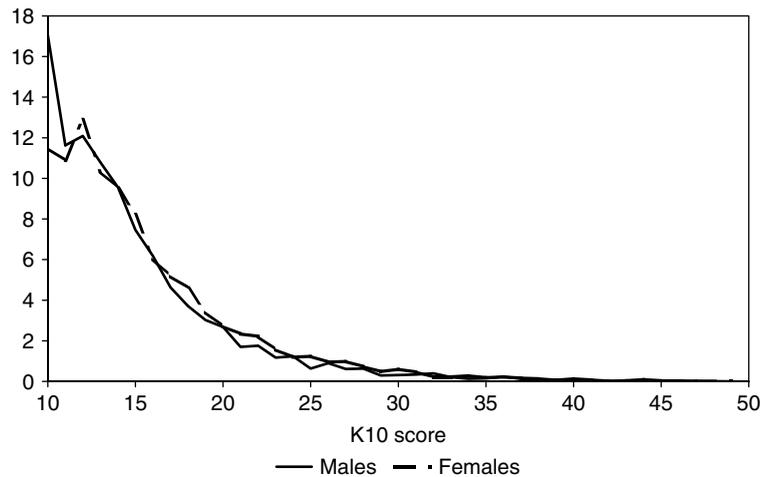Figure 1. Hazard of starting cannabis use by gender (annual percentages)

Figure 2. K10 score (percentages)

20 indicate no mental health problems, and 86% of males and 83% of females fall within this category. Scores in the range of 20–24 indicate a mild mental disorder and 9% of males and 10% of females fall within this category. A moderate mental disorder is indicated by a score between 25 and 29 and a severe mental disorder is indicated by a score of 30 or greater. In this sample, 3% of males and 4% of females suffer a moderate mental disorder and 2% of males and 3% of females suffer a severe mental disorder according to their K10 score.

The distribution of mental health status conditional on cannabis use status (never used, past user, current user) is shown for males and females in Table II. Three main points emerge from examining this table. First, there is a higher prevalence of mental illness among current and past users of cannabis compared to those who have never used. For example, we see that 11% of males

Table II. Mental illness conditional on cannabis use

	Never used	Past user	Current user
(a) *Males*			
No mental health problem	0.89	0.87	0.79
Mental health problem	0.11	0.13	0.21
Mild	0.07	0.08	0.13
Moderate	0.03	0.03	0.04
Severe	0.01	0.02	0.04
(b) *Females*			
No mental health problem	0.86	0.81	0.71
Mental health problem	0.14	0.19	0.29
Mild	0.08	0.11	0.16
Moderate	0.04	0.05	0.08
Severe	0.02	0.03	0.05

and 14% of females who have never used cannabis have a mental health disorder, compared to 21% of males and 29% of females who are current users. Second, past users have a lower prevalence of mental illness than current users. For example, 13% of males and 19% of females who are past users of cannabis are classified as having a mental health disorder, compared to 21% of males and 29% of females who are current users. The third stylized fact that emerges from Table II is that females are more likely than males to suffer mental illness.

The above discussion demonstrates an association between cannabis use and mental health problems that appears to differ between past and current use. The analysis that follows attempts to discern the degree to which this relationship is causal.

3. EMPIRICAL STRATEGY

The decision to start and stop using cannabis as well as the likelihood of experiencing mental illness may be affected by many personal characteristics in addition to circumstances faced in childhood and early adulthood. The most significant challenge posed in investigating potential links between cannabis use and mental health is the impossibility of observing all the personal characteristics and circumstances that might be relevant. According to Pudney (2010), even the most comprehensive longitudinal survey cannot hope to measure every relevant aspect of the individual and his or her environment. Nevertheless, in order to be able to assess whether there is a causal link between cannabis use and mental health common unobserved 'confounding' factors that may be a source of spurious association must be taken into account. It is difficult, if not impossible, to identify characteristics or circumstances that affect cannabis use but not mental health, rendering the use of instrumental variable techniques infeasible.[6] Instead we exploit the discrete-factor approach. This method allows for correlation in unobservables across multiple equations without imposing distributional assumptions and has been used in a wide variety of

---

[6] Note that policy variables relating to cannabis use, such as its legal status and the price of cannabis, are potential candidates for instruments. Given our approach, this would require data on prices and policies from 1966 to the present since our oldest sample members are 50 in 2004 and are assumed to be at risk of uptake from the age of 12. This information is simply not available for cannabis prices. While it may be possible to construct this type of historical series for the legal status of cannabis, it would vary insufficiently for it to be useful since there are only eight states and territories in Australia and the four that have decriminalized the consumption of cannabis have done so rather recently.

applications in health and labor economics (see, for example, Cutler, 1995; Bray, 2005; Van Ours, 2006, 2007; Yang *et al.*, 2009).

The discrete-factor approach was proposed by Heckman and Singer (1984) to address unobserved heterogeneity in hazard rate analysis. It has been further developed by Mroz (1999) for application to regression models with endogenous dummy variables.[7] Mroz demonstrates that when the idiosyncratic error terms for the latent endogenous variable and the outcome of interest have a bivariate normal disturbance, the discrete-factor method compares favorably to the usual maximum likelihood estimator (MLE) in terms of precision and bias. Furthermore, the discrete-factor approximation outperforms both the MLE and the two-stage estimator (TSE) when the disturbances are non-normal.[8]

In our application, we use the discrete-factor approach to account for the unobserved factors affecting the production of mental health, cannabis uptake and quitting in order to obtain reliable estimates of the mental health effects of cannabis use. Identification of this trivariate model with correlated errors comes from functional form assumptions. In the case of cannabis uptake and quitting, we follow Heckman and Singer (1984) and assume mixed proportional hazard functions. Owing to its censored nature, the equation for mental health is based on the Tobit model. Similar to Mroz (1999), identification of unobserved heterogeneity in this equation relies on the linearity of the model for the latent variable and normality of its idiosyncratic error. As with any attempt to discern causal effects of endogenous variables, identification of the parameters of interest is ultimately based on untestable assumptions. We have, however, attempted to explore issues related to identification and model specification in an extensive sensitivity analysis that is reported in Section 4.

Our estimation strategy is implemented in three steps. First, we jointly model the hazard rates for starting and quitting cannabis use. In doing so we pay particular attention to modeling the potentially correlated unobserved heterogeneity driving these processes, which is assumed to come from a discrete distribution representing latent proclivities towards cannabis use. In the second step we model the censored continuous measure for mental health treating cannabis use as exogenous, but accounting for unobserved heterogeneity with respect to susceptibility to mental illness using the discrete-factor approach. In the third step, we marry the bivariate hazard model for cannabis uptake and quitting with the model for mental health, accounting for common unobserved confounding factors by allowing the unobserved discrete factors determining the uptake and quitting of cannabis and the production of mental health to be correlated.

3.1. Dynamics in Cannabis Use

The first part of our econometric strategy focuses on modeling the transitions into and out of cannabis use. We model the rate of uptake of cannabis and the quit rate from cannabis use with a bivariate mixed proportional hazard framework. Concerning the uptake of cannabis we assume that potential exposure to cannabis occurs from the age 12. The starting rate for cannabis use, at time (from age 12) t conditional on observed characteristics x and unobserved characteristics u,

---

[7] In his set-up, Mroz interprets the discrete factors as an unobserved covariate that impacts on the outcome of interest as well as the latent process generating the dummy variable, thereby inducing its endogeneity.

[8] The discrete-factor model also outperforms MLE and TSE in the presence of weak instruments in models with non-normal errors. This is of particular salience given that state-level policy variables are often relied upon to identify the effects of substance use and these policy variables tend to be only weakly predictive of substance use.

is specified as

$$\theta_s(t \mid x, u) = \lambda_s(t) \exp(x'\beta_s + u) \tag{1}$$

where $\lambda_s(t)$ represents individual duration dependence and β_s represents a vector of parameters to be estimated. Unobserved heterogeneity accounts for differences in susceptibility to cannabis. We model duration (age) dependence in a flexible way by using a step function $\lambda_s(t) = \exp(\Sigma_k \lambda_k I_k(t))$, where k ($= 1, \ldots, 15$) is a subscript for age categories and $I_k(t)$ are time-varying dummy variables that are one in subsequent categories. We specify 15 age dummies, 14 of which are for individual ages (age 12, 13, ..., 25) and the last interval is open: ≥ 26 years. Because we also estimate a constant term, we normalize $\lambda_1 = 0$.

The conditional density function for the completed durations until first use can be written as

$$f_s(t \mid x, u) = \theta_s(t \mid x, u) \exp\left(-\int_0^t \theta_s(s \mid x, u)\mathrm{d}s\right) \tag{2}$$

Individuals who initiate cannabis use have a completed duration until first use equal to the age at first use minus 12; individuals who have not used cannabis at the time of the survey have a duration until first use that is right-censored at their current age minus 12.

The quit rate from cannabis use at duration of use τ conditional on observed characteristics x, the age of first use a_f, and unobserved characteristics v is specified as

$$\theta_q(\tau \mid x, a_f, v) = \exp(x'\beta_q + \delta_f a_f + v) \tag{3}$$

where δ_f is a parameter and β_q is a vector of parameters to be estimated. The conditional density function for the completed durations of cannabis use can be written as

$$f_q(\tau \mid x, a_f, v) = \theta_q(\tau \mid x, a_f, v) \exp(-\int_0^\tau \theta_q(r \mid x, a_f, v)\mathrm{d}r \tag{4}$$

As we do not know the actual age at which individuals quit cannabis use, we cannot calculate the exact duration of use and hence cannot estimate the conditional density for completed durations. However, we do know whether or not individuals used cannabis in the 12 months before the survey. For those who had not, we know that their duration of use, τ_s, lies in the interval $[0, a_{s-1} - a_f]$, where a_s is the respondent's age at the time of the survey and a_f is their age of first use. We can therefore account for the uncertainty over the exact age at quitting by integrating the conditional density for durations of use over this range: $\int_0^{\tau_s} f_q\mathrm{d}q = F_q(\tau_s \mid x, a_f, v)$, where F_q is the distribution function of f_q. Individuals who are still using cannabis have a duration of use that is right censored, and for these observations we use the survival function: $1 - F_q(\tau_s \mid x, a_f, v)$.

Modeling the dynamics of cannabis use requires information about characteristics and circumstances faced by individuals at each point in time in which they are confronted with the choice to initiate or quit cannabis use. Information likely to be relevant includes family situation, experiences at school, cannabis supply conditions, the price of cannabis, and the price of other drugs (substitutes and complements). Unfortunately, this type of information is not available in the NDSHS. We note, however, that many of these factors are likely to be endogenous and it is therefore not clear

how one should proceed if they were available.[9] The observable characteristics that we are able to control for are nationality (an indicator for Australian born), whether the respondent identifies themselves as an Aboriginal or Torres Strait Islander, whether they live in a capital city, birth year (to account for birth cohort size effects), the respondent's state of residence at the time of survey and the respondent's education (an indicator for dropping out of school with a 10th-grade education or less) which is used as a proxy for ability.[10] These characteristics are assumed to be known at the time an individual first faces the decision of whether to initiate cannabis use. In the case of the education variable, this requires the assumption that education represents ability and that this ability is known to the individual from the time he first faces the decision to use cannabis. The education variable will not fulfill this requirement if, at the time an individual decides to start using cannabis, he is uncertain as to whether he will drop out of school before completing 10th grade or, if there exist unobserved characteristics that impact both educational attainment and cannabis use (see, for example, Heckman *et al.*, 2006, who allow latent cognitive and non-cognitive skills to determine education and cannabis use). There is also the possibility of reverse causality, in which case cannabis use may result in dropping out of school. The impact of these issues is examined in Section 4.

The potential correlation between the unobserved components in the hazard rates for cannabis uptake and quitting is taken into account by specifying the joint density function for the duration until first use t and the duration of time until quitting use τ conditional on x as

$$h(t, \tau | x, a_f) = \int_u \int_v f_s(t \mid x, u) f_q(\tau \mid x, a_f, v) \mathrm{d}G(u, v) \tag{5}$$

where $G(u,v)$ is a bivariate discrete distribution with n points of support. The probabilities associated with each type are assumed to have a multinomial logit specification: $p_j = \dfrac{\exp(\alpha_j)}{\Sigma_j \exp(\alpha_j)}$, $j = 1, \ldots, n$, with normalization $\alpha_n = 0$.

The parameters of the model are estimated separately for males and females using maximum likelihood and presented in Table III. To select the number of points of support in the joint distribution of discrete factors, an upward-testing approach is used, starting with one point of support and adding additional points of support in a stepwise manner. Beyond four points of support the locations of additional mass points converged to each other and no improvement of the log-likelihood function was found.[11] The points of support reflect the assumption that, conditional on observed characteristics, there exist four distinct 'types' of individuals who are differentiated by their susceptibility to starting and stopping cannabis use.[12] Type 1, represented by (u_1, v_1), has a low starting rate and a positive quit rate; type 2, represented by (u_2, v_2), has

---

[9] We do attempt to address the issue of greater vulnerability early in life in one of our sensitivity analyses in Section 4. Note that variables reflecting the respondent's current circumstance, such as marital status, that are collected as part of the NDSHS are not useful for modeling cannabis uptake because they represent events that may have taken place long after the individual started to use cannabis.

[10] Jacobson (2004) finds a positive correlation between youth cohort size and the prevalence of cannabis use which she largely attributes to a decreased cost of supply due to reduced risk of arrest for selling and informational economies associated with a larger market.

[11] The bottom part of Table III shows the log-likelihood values for the model with 1, 2, and 3 points of support.

[12] According to Gaure *et al.* (2007) it may not be meaningful to interpret the mass-point distribution in terms of representing a corresponding number of distinct types of individuals, as the underlying true heterogeneity distribution may be continuous. However, in the case of cannabis use the interpretation in terms of types is more natural as starting to use cannabis is a discrete choice, as is quitting conditional on use.

Table III. Parameter estimates dynamics in cannabis use

	Males		Females	
	Starting rate	Quit rate	Starting rate	Quit rate
Explanatory variables				
Australian born	0.22 (4.3)**	−0.07 (0.3)	0.45 (7.5)**	0.13 (1.1)
Aboriginal	−0.21 (1.5)	−0.35 (0.4)	−0.24 (1.9)*	0.05 (0.2)
Education ≤10	0.09 (1.7)*	0.38 (1.1)	0.19 (3.7)**	−0.26 (2.3)**
Birth year/10	0.44 (13.8)**	−0.45 (1.7)*	0.78 (22.8)**	0.10 (0.7)
Age of initiation/10	—	0.80 (3.2)**	—	0.61 (5.1)**
Mass points				
u_1, v_1	−6.05 (33.4)**	0.30 (0.3)	−7.12 (30.1)**	−1.27 (2.4)**
u_2, v_2	−3.24 (17.6)**	−∞	−4.33 (17.8)**	−∞
u_3	−∞	—	−∞	—
α_1	1.39 (16.4)**		2.39 (8.4)**	
α_2	−0.24 (2.0)**		−0.24 (0.9)	
α_3	1.42 (18.3)**		2.29 (8.7)**	
Distribution (%)				
$p_1(u_1, v_1)$	40.4		48.4	
$p_2(u_2, v_2)$	7.9		3.5	
$p_3(u_3)$	41.6		43.7	
$p_4(u_1, v_2)$	10.1		4.4	
−Log-likelihood	11,437.6		14,252.2	
−Log-likelihood				
1 mass point	11,796.3		14,524.3	
2 mass points	11,565.8		14,341.4	
3 mass points	11,453.6		14,256.8	

Note: Sample of 4771 males and 6719 females aged 26–50; all estimates include territories fixed effects (7) and a dummy for capital cities both in the starting rates and in the quit rates; note that the starting rates for cannabis use contain 14 age dummies (annually 13–25 and >25 years); absolute *t*-statistics in parentheses; asterisks indicate that the coefficient is different from zero at a ** 5% or * 10% level of significance.

a high starting rate and a zero quit rate; type 3, represented by (u_3), has a zero starting rate (and hence no quit rate); while type 4, represented by (u_1, v_2), has a low starting rate and a zero quit rate. As shown in Table III, the distribution of unobserved heterogeneity implies that 40.4% of the male sample and 48.4% of the female sample fall into the group with a low starting rate and a positive quit rate, 7.9% of males and 3.5% of females belong to the group with a high starting rate and a zero quit rate, while 41.6% of males and 43.7% of females have a zero starting rate. Finally, 10.1% of males and 4.4% of females have a low starting rate and a zero quit rate.

The results from estimation indicate that Australian-born males and females with a low level of education have a higher uptake of cannabis than foreign-born individuals or those with greater than a 10th-grade education. Those born in more recent years have a higher starting rate than those born earlier, and Aboriginal females are less likely to initiate into cannabis use than other Australian females. In terms of quit rates, we find that females with a low level of education are less likely to quit compared to those females with more than a 10th-grade education. In line with Van Ours and Williams (2007), we also find that age of initiation has a large effect on the quit rate, with those initiating into cannabis use early in life having a lower quit rate.

3.2. Determinants of Mental Health Assuming Exogenous Cannabis Use

Figure 2 shows that, in our sample, the distribution of K10 scores has a significant proportion of observations at the lower limit value of 10. It also shows that the distribution of K10 scores is skewed to the right. To account for these features, we model the production of mental health using a Tobit model for log (K10). The natural log of latent mental health of individual i, m_i^*, is assumed to depend upon a vector of observed characteristics x_m, past and current cannabis use, and unobserved characteristics μ:

$$m_i^* = x_m'\beta_m + \delta_1 cc_i + \delta_2 cp_i + \mu_i + \varepsilon_i \qquad (6)$$

where cc_i is a dummy variable indicating whether or not an individual is a current user of cannabis (has used in the past 12 months), cp_i is a dummy variable indicating whether an individual is a past user of cannabis (has used in lifetime but not in the past 12 months), and the parameters δ_1 and δ_2 and the vector of parameters β_m are to be estimated. Observed mental health, m_i, is measured by the natural log of the K10 score.[13] Given the censoring point of 10 for the K10, the relationship between latent mental health m_i^* and observed mental health m_i is given by $m_i = \log(10)$ if $m_i^* \leq \log(10)$ and $m_i = m_i^*$ if $m_i^* > \log(10)$. The main parameters of interest are δ_1 and δ_2, which measure the effect of current and past cannabis use on mental health, respectively. The observed characteristics, x_m, include the respondent's age, their nationality (an indicator for Australian born), whether the respondent identifies themselves as an Aboriginal or Torres Strait Islander, whether they live in a capital city, the respondent's state of residence at the time of survey, the respondent's education (an indicator for dropping out of school with a 10th-grade education or less) and their marital status (indicators for married and divorced with single or widowed as the comparison category). As with the dynamics of cannabis use, there exists the potential for unobserved characteristics that determine mental health to also determine education and marital status. The sensitivity of our results to this issue is explored below.

The error term for individual i is assumed to be composed of two components. The first is a a discrete factor μ_i, which is intended to capture unobserved susceptibility to mental illness. The second component of the error term is drawn from a normal distribution, $\varepsilon_i \sim N(0, \sigma^2)$. Since we do not have panel data, the identification of the discrete factor relies on the assumption that ε_i is normally distributed.

As before, we assume that the probabilities associated with the discrete factors follow a logit specification. Table IV contains the maximum likelihood estimates of the relevant parameters. We found no improvement in the likelihood function beyond two mass points, implying two distinct 'types' in our sample.[14] The first mass point is estimated to be 2.512 for males and 2.632 for females. This implies K10 sores of 12.3 (exp(2.512)) and 13.9 (exp(2.632)), respectively. Since a K10 score of less than 20 is indicative of an absence of mental illness,

---

[13] A priori, it is not clear whether to use a linear or log-linear specification for the K10 score. We used the pseudo-R^2 as guidance. The pseudo-R^2 is calculated as $(1 - \frac{L_u}{L_0})$, where L_u is the value of the log-likelihood of the full model and L_0 is the value of the log-likelihood of the model with only an intercept. For the logarithmic specification we found a pseudo-R^2 of 0.062 both for males and females; for the linear specification we found a pseudo-R^2 of 0.007 for males and 0.008 for females.

[14] Note that the explanatory variables are specified as deviations from the mean. Also note that the bottom part of Table IV shows that the introduction of two mass points has a big effect on the value of the log-likelihood, while the estimated effects of cannabis use on mental health are very similar to a specification without mass points.

Table IV. Parameter estimates of mental health: log K10

	Males	Females
(a) 2 mass points		
Explanatory variables		
Age/10	−0.027 (3.9)**	−0.020 (3.6)**
Australian born	−0.013 (1.2)	−0.004 (0.5)
Aboriginal	0.107 (3.4)**	0.062 (2.4)**
Married	−0.062 (5.5)**	−0.062 (6.1)**
Divorced	0.052 (3.2)**	0.030 (2.0)**
Education ≤ 10	0.009 (0.7)	−0.003 (0.4)
Effect of cannabis use		
Current user	0.120 (9.8)**	0.131 (10.6)**
Past user	0.059 (5.8)**	0.067 (8.3)**
σ	0.264 (47.1)**	0.249 (64.7)**
Mass points		
μ_1	2.512 (276.8)**	2.632 (119.1)**
μ_2	3.089 (101.4)**	3.203 (120.1)**
α_1	2.125 (13.8)**	1.830 (20.9)**
Distribution (%)		
p_1	89.3	86.2
p_2	9.7	13.8
−Log-likelihood	2005.9	2485.9
(b) 1 mass point		
Effect of cannabis use		
Current user	0.122 (9.2)**	0.140 (10.2)**
Past user	0.057 (5.1)**	0.069 (7.8)**
−Log-likelihood	2059.9	2618.4

Note: Sample of 4771 males and 6719 females; in all estimates the explanatory variables are specified as deviations from their mean value and they include territories fixed effects (7) and a dummy for capital cities; asterisks indicate that the coefficient is different from zero at a ** 5% or * 10% level of significance.

this group is considered to have a low susceptibility to mental health problems. The second mass point is 3.089 for males and 3.203 for females, which translates into K10s of 22.0 and 24.6. Given that a K10 score between 20 and 25 identifies a person as having mild mental health problem, those belonging to this group are considered to be susceptible to mental health problems.

Table IV shows that, all else being equal, married respondents are in better mental health than those who are single or widowed and that divorcees are in worse mental health than those who are single or widowed. Older respondents are, on average, in better mental health than their younger counterparts, and Aboriginals are in worse mental health than non-Aboriginals. We find no evidence that a low level of education is associated with worse mental health.

The main parameters of interest are those measuring the effects of current and past cannabis use on mental health. As shown in Table IV, we find that both men and women who currently use cannabis have a higher K10 score than those who have never used cannabis, where the increase in K10 attributable to current cannabis use is similar in size and magnitude to the effect of a married person becoming divorced. We also find that past users of cannabis have a higher K10 score compared to their counterparts who have never used. The increase in the K10 score attributable to past use is around half of that associated with current use.

3.3. A Joint Model of Cannabis Use and Mental Health

In estimating the joint model for cannabis dynamics and the production of mental health, we start by assuming that the distribution of unobserved heterogeneity has eight points of support, reflecting two types in terms of susceptibility to mental illness combined with four types in terms of susceptibility to cannabis use. As before, the associated probabilities are assumed to have a multinomial logit specification: $p_j = \frac{\exp(\alpha_j)}{\Sigma_j \exp(\alpha_j)}$, $j = 1, \ldots, 8$, with the normalization $\alpha_8 = 0$.

The three-equation system is estimated using maximum likelihood and the relevant estimates are presented in the first panel of Table V. We do indeed find that the joint distribution of unobserved heterogeneity for the cannabis starting and quit rates and the production of mental health has eight points of support. The vast majority of the sample belong to two groups, each of which has a low susceptibility for mental health problems. For example, 36.9% of males and 41.4% of females belong to the group characterized by a low positive cannabis starting rate, a positive cannabis quit rate and low susceptibility to mental health problems (type 1), and 36.9% of males and 38.1% of females have a zero cannabis starting rate and a low susceptibility to mental illness (type 3). While this distribution demonstrates a correlation between cannabis use and susceptibility to mental illness, the correlation does not appear to be very strong. At the 10% level of significance, a likelihood ratio (LR) test of the null hypothesis of no correlation is rejected for men, confirming that common unobserved factors is an issue for this sample. For women this is not the case.

A comparison of the results from estimating models with and without accounting for common unobserved factors (comparing the first panel of Table V and the parameter estimates in Table IV) reveals that the causal effect of cannabis use on mental health is overestimated if one does not account for these unobserved common factors. This implies that those who are more likely to start using cannabis are also more likely to have mental health problems. However, as shown, the parameter estimates representing the effect of current and past cannabis use on mental health are not much affected by whether or not correlation in unobserved characteristics is accounted for. From this we conclude that after accounting for the potentially confounding effect of unobserved characteristics cannabis use has an adverse impact on mental health. Moreover, while this effect is greater for current use, it persists well after use ceases.

4. SENSITIVITY ANALYSIS AND SIMULATIONS

We next investigate the sensitivity of our findings to assumptions made in modeling the relationship between cannabis use dynamics and mental health. The results from doing so are summarized in Table V.

First, we investigate the impact of recall error by estimating our model over the subsample of 26- to 35-year-olds. To the extent that respondents make mistakes in the age they report first using cannabis, our parameter estimates of the impact of cannabis use on mental health based on the full sample and reported in Table V Panel 1 are likely to be biased towards zero. We expect that recall error is less of a problem among the 26–35 age group (as less calendar time has elapsed since they first used and perhaps quit cannabis) and hence results based on these data should be more reliable. The results for this younger subsample are reported in Table V Panel 2. A comparison of results based on estimation over the full sample with those based on the younger sample reveals no significant differences, suggesting that recall error is not an important source of bias in the data used in this analysis.

Table V. Effects of cannabis use on log K10 score

	Males		Females	
1. Baseline estimate	α	p (%)	α	p (%)
(u_1, v_1, μ_1)	3.202 (13.6)**	36.9	3.972 (10.4)**	41.4
(u_2, v_2, μ_1)	1.479 (5.7)**	6.6	1.208 (3.1)**	2.6
(u_3, μ_1)	3.203 (13.6)**	36.9	3.890 (10.5)**	38.1
(u_1, v_2, μ_1)	1.741 (6.7)**	8.6	1.589 (4.3)**	3.8
(u_1, v_1, μ_2)	0.865 (3.3)**	3.6	2.147 (5.4)**	6.7
(u_2, v_2, μ_2)	−0.191 (0.6)	1.2	0.047 (0.1)	0.8
(u_3, μ_2)	1.168 (4.8)**	4.8	2.011 (5.2)**	5.8
(u_1, v_2, μ_2)	—	1.5	—	0.8
Cannabis: current user	0.107 (6.8)**		0.112 (6.8)**	
Cannabis: past user	0.070 (5.6)**		0.065 (6.6)**	
LR test selectivity	7.4*		5.6	
2. Age ≤35				
Cannabis: current user	0.112 (5.3)**		0.097 (4.0)**	
Cannabis: past user	0.052 (2.6)**		0.063 (3.8)**	
3. By age of initiation				
≤15				
Cannabis: current user	0.105 (4.4)**		0.132 (4.8)**	
Cannabis: past user	0.101 (4.8)**		0.119 (6.8)**	
>15				
Cannabis: current user	0.109 (6.3)**		0.103 (5.6)**	
Cannabis: past user	0.066 (5.1)**		0.057 (5.6)**	
LR test age specificity	2.8		15.0**	
4. Intensity of cannabis use				
Every day	0.151 (6.5)**		0.191 (6.2)**	
Once a week or more	0.143 (6.0)**		0.131 (4.7)**	
About once a month	0.080 (2.4)**		0.115 (3.4)**	
Every few months	0.060 (2.1)**		0.109 (3.3)**	
Once or twice a year	0.066 (2.7)**		0.078 (3.3)**	
Past user	0.069 (5.6)**		0.065 (6.6)**	
LR test intensity matters	13.6**		8.4*	

Note: The parameter estimates for the starting rates, quit rates and the other parameter estimates for mental health are not reported because they are similar to those in Tables III and IV; all estimates contain 8 mass points except those for young females (age ≤35), where we could only identify 6 mass points. Sample of 4771 males and 6719 females except for estimate 2: 1791 males and 2756 females; absolute *t*-statistics in parentheses; asterisks indicate that the coefficient is different from zero at ** 5% or * 10% level of significance.

Second, we examine whether the impact of cannabis use on mental health varies by age of first use. Specifically, we allow the effects of past and current use on mental health to differ for those who first used cannabis before the age of 16 and those who were 16 or older when they first tried cannabis. An LR test is used to test the hypothesis that the effects of cannabis use differ by age of first use. As shown in the third panel of Table V, we find significant differences in the impact of cannabis use on mental health by the age of uptake for females but not males. Specifically, we find that current and past use of cannabis produces significantly greater increases in the K10 score of females who first used cannabis before the age of 16 compared to those who first used at 16 years or older.

We also investigate whether the mental health effects of current cannabis use depend upon the frequency of use. To do so, we measured frequency of use in the past year with a set of indicators for the categories use every day, use weekly, use monthly, use every few months, and

use once or twice a year, in addition to the categories past use and never use. As shown, the mental health effects of cannabis use increase with the frequency of use. For example, on average males who use once or twice a year have a K10 score which is similar to males who stopped using cannabis, whereas males who use daily have a K10 score that is approximately 15.1% larger than a comparable male who has never used. For females, we find similar results.

In addition to the sensitivity analyses reported in Table V, we also investigated the robustness of our results to several other aspects of model specification. First, we examined the impact of the potential endogeneity of marital status and education on the estimated effect of cannabis use on mental health. We did this by comparing results from our baseline model (Table V Panel 1) with estimates from a model which omitted education and marital status from all three equations. The estimated effects of cannabis use (past and current) are quite robust, suggesting this potential source of model misspecification is not driving our findings. Second, we allowed for more flexible cohort effects by replacing the continuous year of birth variable with indicators for birth cohort (born in the 1950s, the 1960s and the 1970s). The results from this model are almost identical to those from the original specification. We also explored the potential for heterogeneity in treatment effects by allowing all parameter estimates in the mental health equation to differ by whether the individual was a never user, current user or past user of cannabis. On the basis of an LR test we are unable to reject the null hypothesis of homogeneity in the effects of cannabis use on mental health. Finally, we considered the sensitivity of our results to specifying mental health as a continuous (censored) variable. As an alternative approach we constructed an ordinal categorical variable for mental health (no mental health problems, mild mental health problems, moderate mental health problems and severe problems). The estimated effects from ordered probit models of mental health are very similar to our baseline estimates in Table V Panel 1.

In order to illustrate the impact of cannabis use on mental health as measured by the K10 score, we use the parameter estimates contained in Table V Panel 4 to perform simulations. We present scenarios in which individuals vary in their frequency of cannabis use, ranging from never users to daily cannabis users. All simulations are done separately for males and females. The results are presented in Table VI.[15] The simulations show that, for males, the impact of using cannabis weekly or more often produces a substantial increase in the prevalence of mental illness compared to those who have never used, used in the past, and those who use less frequently. The prevalence of mental illness among frequent cannabis-using males (using weekly or more often) is predicted to be in the range of 14.3–14.8%. In comparison, otherwise similar males who have never used cannabis are predicted to have a prevalence of mental illness of 8.0%. This increase occurs predominantly among the categories of mild and severe mental illness. For example, Table VI shows that the prevalence of severe mental illness in males who use cannabis weekly (or daily) is double that of males who have never used cannabis, at 2.3% (or 2.4%) compared to 0.9%. The simulations also show that less frequent use of cannabis produces a small increase in the prevalence of mental health problems, similar in magnitude to the increase associated with past cannabis use. For example, the prevalence of severe mental distress among males who use cannabis once a month is 1.5%, compared to 1.4% among past users. Intensity of use is found to have a similar effect on females, except that using every few months rather than weekly appears to be the critical use category and females generally experience a higher rate of mental health problems compared to males.

---

[15] The simulations assume that individuals are married, Australian born, non-Aboriginal aged 40 years old, who have more than 10 years of education.

Table VI. Results of simulation effects of cannabis use on K10 score; different types of individuals

	Never used	Past user	Uses once or twice a year	Uses every few months	Uses once a month	Uses weekly	Uses daily
Males							
No problem	92.0	84.9	89.5	89.7	88.9	85.7	85.2
Mild problem	5.2	6.6	6.5	6.5	6.9	8.6	8.9
Moderate problem	1.9	2.6	2.6	2.4	2.7	3.4	3.5
Severe problem	0.9	1.4	1.4	1.4	1.5	2.3	2.4
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Females							
No problem	90.3	87.6	87.0	85.4	85.2	84.2	80.3
Mild problem	6.2	7.5	7.8	8.6	8.6	9.2	11.2
Moderate problem	2.4	3.2	3.3	3.7	3.8	4.0	4.8
Severe problem	1.1	1.7	1.9	2.3	2.4	2.6	3.7
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Note: Individual: Australian born, non-Aboriginal married individual age 40 having an education of more than 10 years.

5. CONCLUSIONS

Despite the widely held belief that cannabis is a benign drug, there is increasing evidence of an association between its use and mental illness. Our paper investigates the extent to which this association is causal. In order to do so, we account for the potential for common unobserved factors affecting cannabis use and mental health. We do this by modeling cannabis uptake and quitting jointly with the production function for mental health, allowing the unobserved heterogeneity determining each to be correlated. Our results suggest that cannabis use does have an adverse effect on mental health, with frequent current use having a larger effect than infrequent current use or past use. We also find that unobserved factors that make individuals more susceptible to cannabis use also make them more susceptible to mental illness. While accounting for common unobserved factors reduces the size of the estimated effects of current and past cannabis use on mental health, it does not eliminate them.

It should be reiterated that our findings are subject to caveats related to limitations of our data and assumptions underlying our identification strategy. For example, due to the absence of information on the age of onset of mental health problems, this study has focused on common unobserved factors as the source of the endogenity of cannabis use in the production of mental health. Reverse causality remains a potential issue, although previous studies suggest that its effects are likely to be small. A second shortcoming of the data is that the information on age at first use of cannabis is retrospectively reported and potentially subject to reporting errors. We attempt to investigate this issue as part of our sensitivity analysis and find no evidence that, for our sample, this is an important source of bias in our estimates. As part of our extensive sensitivity analysis, we also investigated the impact of modeling assumptions including the functional form of the mental health equation and homogeneity of treatment effects. Overall, we found our results to be very robust and, within the limitation of this type of research, indicate that cannabis use leads to worse mental health. We find that frequent cannabis use poses the greatest risk to mental health. This has wide-ranging implications not just for the individuals (and their families) but also for the wider society. Mental illness imposes sizable costs on the economy, the most significant of which is attributable to loss in productivity. While these results are striking, it is nonetheless important

to emphasize that only frequent cannabis use is found to cause large increases in the likelihood of mental illness, and only a fraction of cannabis users fall into this category.

REFERENCES

Andreasson S, Allebeck P, Engstrom A, Rydberg U. 1987. Cannabis and schizophrenia: a longitudinal study of Swedish conscripts. *Lancet* **2**: 1483–1486.

Andrews G, Slade T. 2001. Interpreting scores on the Kessler Psychological Distress Scale (K10). *Australian and New Zealand Journal of Public Health* **25**: 494–497.

Arseneault L, Cannon M, Witton J, Murray RM. 2004. Causal association between cannabis and psychosis: examination of the evidence. *British Journal of Psychiatry* **184**: 110–117.

Australian Bureau of Statistics. 2008. *National Survey of Mental Health and Wellbeing: Summary of Results, 2007. ABS catalogue* 4326. Australian Government Publishing Service: Canberra.

Australian Institute of Health and Welfare and the Australian Government Department of Health and Ageing. 2005. *National Drug Strategy Household Survey, 2004* [computer file]. Australian Social Science Data Archives, The Australian National University: Canberra.

Boys A, Farrell M, Taylor C, Marseden J, Goodman K, Brugha T, Bebbington P, Jenkins R, Meltzer H. 2003. Psychiatric morbidity and substance use in young people aged 13–15 years: results from the Child and Adolescent Survey of Mental Health. *British Journal of Psychiatry* **182**: 509–517.

Bray JW. 2005. Alcohol use, human capital, and wages. *Journal of Labor Economics* **23**: 279–312.

Cutler D. 1995. The incidence of adverse medical outcomes under prospective payment. *Econometrica* **63**: 29–50.

Degenhardt L, Hall W, Lynskey M. 2003. Exploring the association between cannabis use and depression. *Addiction* **98**: 1493–1504.

Fergusson DM, Horwood J. 1997. Early onset cannabis use and psychological adjustment in young adults. *Addiction* **92**: 279–296.

Fergusson DM, Horwood J, Swain-Campbell N. 2002. Cannabis use and psychosocial adjustment in young adults. *Addiction* **97**: 1123–1135.

Fergusson DM, Horwood J, Ridder E. 2005. Tests of causal linkages between cannabis use and psychotic symptoms. *Addiction* **100**: 354–366.

Gaure S, Røed K, Zhang T. 2007. Time and causality: a Monte Carlo assessment of the timing-of-events approach. *Journal of Econometrics* **141**: 1159–1195.

Hall W. 2006. Cannabis use and the mental health of young people. *Royal Australian and New Zealand Journal of Psychiatry* **40**: 105–113.

Heckman JJ, Singer B. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* **52**: 271–320.

Heckman JJ, Stixrud J, Urzua S. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* **24**: 411–482.

Henquet C, Krabbendam L, Spauwen J, Kaplan C, Lieb R, Wittchen H, van Os J. 2005. Prospective cohort study of cannabis use, predisposition for psychosis, and psychotic symptoms in young people. *British Medical Journal* **330**: 11–15.

Jacobson M. 2004. Baby booms and drug busts: trends in youth drug use in the United States, 1975–2000. *Quarterly Journal of Economics* **119**: 1481–1512.

Johnston LD, O'Malley PM, Bachman JG, Schulenberg JE. 2006. Teen drug use continues down in 2006, particularly among older teens. University of Michigan New and Information Services: Ann Arbor, MI. Available: www.monitoringthefuture.org [18 August 2007].

Kalant H. 2004. Adverse effects of cannabis on health: an update of the literature since 1996. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **28**: 849–863.

Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, Walters EE, Zaslavsky AM. 2002. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine* **32**: 959–976.

Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, Howes MJ, Normand S-LT, Manderscheid RW, Walters EE, Zaslavsky AM. 2003. Screening for serious mental illness in the general population. *Archives of General Psychiatry* **60**: 184–189.

Macleod J, Oakes R, Copello A, Crome I, Egger M, Hickman M, Oppenkowski T, Stokes-Lampard H, Smith GD. 2004. Psychological and social sequelae of cannabis and other illicit drug use by young people: a systematic review of longitudinal, general population studies. *Lancet* **363**: 1579–1588.

McGee R, Williams S, Poulton R, Moffitt T. 2000. A longitudinal study of cannabis use and mental health from adolescence to early adulthood. *Addiction* **95**: 419–503.

Mroz TA. 1999. Discrete factor approximations in simultaneous equation models: estimating the impact of a dummy endogenous variable on a continuous outcome. *Journal of Econometrics* **92**: 233–274.

Patton GC, Coffey C, Carlin JB. 2002. Cannabis use and mental health in younger people: cohort study. *British Medical Journal* **325**: 1195–1198.

Pirkis J, Burgess P, Kirk P, Dodson S, Coombs T. 2005. *Review of standardized measures used in the National Outcomes and Casemix Collection (NOCC)*. Australian Mental Health Outcomes and Classification Network: Melbourne/Brisbane/Sydney.

Pudney S. 2010. Drugs policy: what should we do about cannabis? *Economic Policy* **61**: 165–211.

Rey JM, Sawyer MG, Raphael B. 2002. Mental Health of teenagers who use cannabis: results of an Australian survey. *British Journal of Psychiatry* **180**: 216–221.

Van Os J, Bak M, Hanssen M, Bijil RV, Graaf ER, Verdoux H. 2002. Cannabis use and psychosis: a longitudinal population based study. *American Journal of Epidemiology* **156**: 319–327.

Van Ours JC. 2006. Cannabis, cocaine and jobs. *Journal of Applied Econometrics* **21**: 897–917.

Van Ours JC. 2007. The effects of cannabis use on wages of prime age males. *Oxford Bulletin of Economics and Statistics* **69**: 619–634.

Van Ours JC, Williams J. 2007. Cannabis prices and dynamics of cannabis use. *Journal of Health Economics* **26**: 578–596.

Van Ours JC, Williams J. 2009. Why parents worry: initiation into cannabis use by youth and their educational attainment. *Journal of Health Economics* **28**: 132–142.

White V, Hayman J. 2006. *Australian secondary school students' use of over-the-counter and illicit substances in 2005*. Report prepared for Australian Government Department of Health and Ageing, Centre for Behavioural Research in Cancer, Cancer Control Research Institute, Cancer Council Victoria.

Williams J, Skeels CL. 2006. The impact of cannabis use on health. *De Economist* **154**: 517–546.

World Health Organization. 2003. *Investing In Mental Health*. Department of Mental Health and Substance Dependence, Noncommunicable Diseases and Mental Health, WHO: Geneva.

Yang Z, Gilleskie D, Norton E. 2009. Health insurance, medical care and health outcomes: a model of elderly dynamics. *Journal of Human Resources* **44**: 47–114.