ECONOMIC POLICY

EUROPEAN COMMISSION

Economic Policy
Forty-Ninth Panel Meeting

Supported by the European Commission
Brussels 24-25 April 2009

Drugs Policy – What Should We Do About Cannabis?

Stephen Pudney

The organisers would like to thank the European Commission for their support.

The views expressed in this paper are those of the author(s) and not those of the funding organization(s).

Centre for Economic Policy Research

CES

PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

Drugs policy - what should we do about cannabis?

Stephen Pudney

Institute for Social and Economic Research, University of Essex

1. INTRODUCTION

Almost any consumption good can be harmful. King Henry I died from a surfeit of lampreys. We eat too much and suffer obesity-related disease, drive around in cars and die from accidents and lack of exercise, drink excessive alcohol and risk alcoholism and liver damage, smoke tobacco and contract lung cancer, use cannabis, opiates, cocaine and amphetamines and risk dependence and impaired cognitive function. Many forms of consumption are controlled to varying degrees, with policy instruments ranging from health education campaigns, advertising controls, indirect taxation, through controls on product quality, supply restrictions and local consumption embargoes, to complete bans on supply and purchase, backed by legal penalties. Complementary to this system of primary policy is the health and social care system which may offer treatment for the adverse effects of consumption and services to combat dependence.

It is sometimes hard to see much evidence of rationality in this area of policy-making and the policy debate is dominated by strong views on either side, often with little theoretical or empirical basis. The public demand for policy action is as interesting a subject for study as the demand for drugs itself.

In this paper, we concentrate on one particular substance, cannabis (or marijuana, hashish, etc.), which derives from the hemp plant *cannabis sativa*, particularly subspecies *indica*. Products of the cannabis plant are commonly classified into three main forms: cannabis resin and other extracts (hashish or hash oil); 'regular' herbal

I am grateful to Priscila Ferreira for research assistance. This work was supported by the Economic and Social Research Council through the MiSoC research centre (award no. RES518285001).

cannabis (marijuana); and preparations of the female flowering tops (sinsemilla, skunk). The primary psycho-active constituent is Δ^9-tetrahydrocannabinol (THC) but cannabis has a complex chemical structure and there are many other chemical components, whose physical and psychological impacts are not well understood. One component, cannabidiol (CBD), may have a significant antipsychotic effect, so that the mental health consequences of cannabis use may depend on the relative THC and CBD content (Zuardi et al. 1982).

Cannabis is easily cultivated and widely available in most developed countries, although there are striking international differences in prevalence levels and trends. It is an illegal substance everywhere, although with considerable variation in the rigour of enforcement and severity of punishment. As a consequence of its illegality, consumption measurement is difficult, particularly for the purposes of international comparison. Arguably the most reliable source of international comparative data is the ESPAD group of school surveys which use harmonised questions on comparable samples of 16-year old schoolchildren in a large number of European countries, and the similar US *Monitoring the Future* survey. Figure 1 summarises the proportion of survey respondents who reported having used cannabis at any time in the past in 1995, 1999 and 2003. Over that 8-year period, prevalence increased in all the former socialist countries and also in Finland, Italy and Portugal, with little obvious trend elsewhere. Self-reported prevalence in these school cohorts is remarkably high – exceeding 25% in eight of the twenty countries.

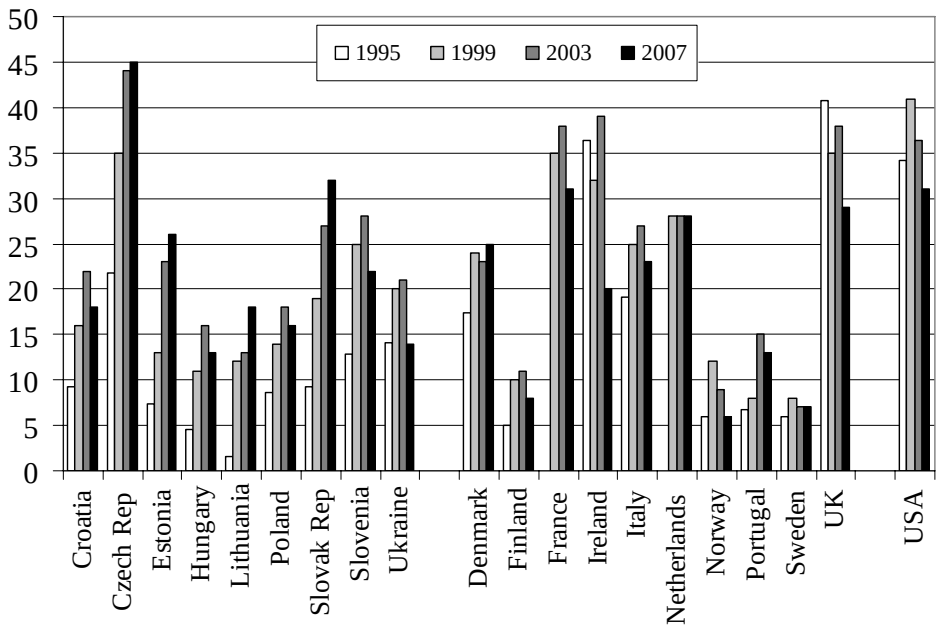


Figure 1. Lifetime cannabis prevalence among 16-year old school students in Europe and the USA (ESPAD school surveys and US Monitoring the Future survey)

Cannabis use tends to start early or not at all. The left-hand panel of Figure 2, based on 2003 cross-section survey data for England and Wales shows the typical pattern of (current) cannabis use by age: a rapid rise from the early teens to a peak in the mid to late teens, followed by a slower decline, with prevalence falling to near zero by fifty. The right-hand panel shows lifetime prevalence, indicating that over half the people interviewed had used cannabis by their early twenties. The decline following that peak is a reflection of the rising macro trend in cannabis consumption over the previous three decades. Few people commence cannabis use over the age of twenty-five, so there is a strong cohort effect: many people in earlier generations were never initiated into drug use. This in turn means that the future decline in cannabis consumption by people currently in their twenties is unlikely to be as rapid as the left-hand panel of Figure 2 suggests. The age profile in Figure 2 is generally representative of the pattern observed in developed countries, except that onset is rather earlier, and peak prevalence higher, in the UK than most other countries.

Figures 1 and 2 reveal a relatively low age of initiation into cannabis use for a substantial minority of users and this is a particular concern, given the strong empirical association between early first use and the length and intensity of the period of subsequent cannabis consumption (Pudney 2004).

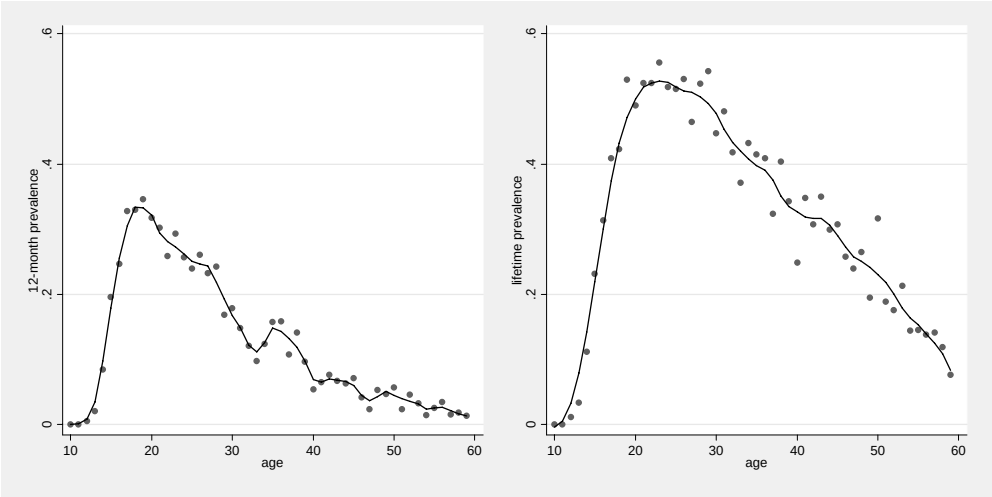


Figure 2. The age profile of cannabis prevalence (i) within the previous 12 months (ii) any time in the past (England and Wales Offending Crime and Justice Survey 2003)

Prices are more difficult to measure than consumption and most available price data are a by-product of criminal intelligence and market disruption activity by enforcement agencies. Consequently, they fall short of the random sampling of price quotations for standard commodity underlying other price indexes. Prices derived from surveys of drug users often contain large outliers, in part because, in the absence of a standard marketed product, there is considerable confusion about the physical units involved. However, the

general picture in recent years is one of real price stability or decrease at the retail level. For example, in the US, the average cannabis price is estimated to have halved in real terms from 1991 to 1998 and to have risen only slightly from 1999 to 2003 (Caulkins et al 2004). In the UK, price figures gathered by the criminal intelligence service suggest a fall of around 15% in nominal terms over 2000-2004 (Pudney et al. 2006), while average prices recorded in an independent internet survey (IDMU 2008) suggest a slight fall in nominal terms over 2002-2006. The picture is similar in Australia, where a 35% fall in the real cannabis price has been reported over the 1990s, mostly occurring in the first couple of years of the decade (Clements 2002). It is clear that policy has largely failed to bring about a rise in the real price of cannabis, which remains one of the cheapest ways to achieve a 'high' – generally much cheaper than alcohol (Kleiman 1992).

Given the failure of current policy on cannabis to prevent large-scale consumption, what, if anything, should we be doing to change the policy environment? No serious commentator disputes the fact that cannabis consumption is harmful in various ways, but this is not in itself sufficient justification for a prohibitionist policy. In this paper, we first set out the arguments for policy intervention in the cannabis market, then consider some directions of policy change that have been called for. We then consider the research evidence available as a basis for policy review, concluding that it is hard to find any strong theoretical argument or empirical evidence to support the current regime of prohibition. We argue that there is no strong reason to prefer prohibition to the alternative of legalisation of cannabis with harms controlled by regulation and taxation. Given this conclusion and the much wider prevalence of cannabis than of harder drugs, we suggest that there is no strong case for covering cannabis by the same prohibition that applies to cocaine and opiates. A reasonable way forward is to remove cannabis production and consumption (but *not* trade) from the current prohibitionist UN drug control treaties, to allow individual countries to adopt their own policies.

The most fundamental problem for research is the difficulty of understanding causal influences on cannabis consumption, especially the link between cannabis consumption and subsequent harms to consumers, purely from observation of behaviour within the current prohibitionist policy environment. In section 3.4 we contribute new evidence on this issue. Removal of international prohibition would make possible a much wider range of policy initiatives, generating new 'quasi-experimental' international evidence, including full legalisation with regulated supply and taxation.

2. CANNABIS HARMS AND ECONOMIC ARGUMENTS FOR INTERVENTION

It is useful to distinguish four distinct bases for a given legal position on cannabis: First is the orthodox idea of *imperative* law as an externally-imposed obligation with enforcement and legal sanctions imposing costs on wrongdoers, who are thus presented with a disincentive to use the drug. Second, law may be seen as *education*, with the illegal status of cannabis used to signal the existence of harms to ill-informed (and hard-to-inform) citizens. Third, the theory of *expressive law* (Cooter 1998) sees policy instead

as a device for signalling a moral position or social norm, which may influence behaviour either directly, by inducing individuals to internalise that norm thus effecting a change in preferences; or indirectly, by highlighting a particular social equilibrium and influencing equilibrium selection. There is some evidence in favour this expressive role in simple experimental games (Bohnet and Cooter 2001, Galbiati and Vertova 2008), where drawing attention to a particular equilibrium without altering incentives has been found to influence outcomes. Related to this, a fourth role of legislation is as a public *consumption good*. Supporting a law prohibiting specific behaviour can be used as means of distinguishing oneself from groups which happen to be characterised by that behaviour, even if the behaviour in question is not especially harmful. Even if ineffective, the law can generate utility for members of the pro-legislation group by reinforcing a positive self-image.

Given these various roles, how should we view the law on cannabis? The education and expressive roles of cannabis legislation are often emphasised but, young people (particularly those with developmental problems making them vulnerable to risky behaviour) are not well represented in, or closely identified with, the institutions of legislative power. For example, Torry-Purta et al (2001) analyse civic trust among 14 year-olds in 28 countries, finding that only 20-25% "always trust" the courts and police, around 10% national and local government and only 4% political parties. See also Smith (1999). It is plausible to argue that society is polarised to some degree, with parallel and antithetic social norms operating in different social groups – particularly so in relation to cannabis, which has a historical association with a counter-culture.[1] Prohibition as expressive law may strengthen adherence to the non-consumption social norm among the 'establishment' group, whilst simultaneously strengthening the dissident identity of the counter-culture group who are at highest risk of starting cannabis use.

Historical anecdote tends to support this idea of polarisation and the interpretation of legislation as a public consumption good, particularly given the apparently arbitrary nature of the set of substances identified as the objects of drug policy. In Hogath's London, gin was the subject of a moral panic leading to calls for tough policy that were met with the introduction of a series of stringent laws (Warner 2003). Today, gin - preferably with tonic water (which was originally produced as a drug) - is a highly respectable drink consumed by many who deplore cannabis consumption and call for tough drugs policy. The increasing disdain for opiates in Britain at the end of the 19th century stemmed in part from a presumed association of opiates with the recent influx of Chinese immigrants: a view which conveniently overlooked Britain's role in promoting the opium trade. The more recent heated debate over cannabis began as a reaction to the counter-culture of the 1960s and has also sometimes been linked to particular minority ethnic groups. If this interpretation is correct, it suggests that the expressive power of the establishment group's support for prohibition is unlikely to be great. We present evidence consistent with this in section 3.3. Perhaps the most powerful objection to the view of prohibition as effective education or enforcement of a social norm is its

[1] This is similar to the point made by Kuran (1998) in the context of ethnic identity.

historical lack of success: cannabis was illegal throughout the 1970s and 1980s when cannabis prevalence rose strongly in most developed countries

The case for intervention in the cannabis market hinges on five main issues. First, the direct internal harms experienced by users themselves and the extent to which the risks are fully considered. This in turn depends on the information available to potential cannabis users and their capacity to make 'rational' decisions. Second, the direct external costs of cannabis use incurred by others besides the user. Third, the indirect or knock-on costs stemming from the additional engagement in further illicit drug use that might be induced by experience of cannabis. Fourth, the potential effectiveness of available anti-cannabis policy instruments; and fifth, the costs of implementing those policies. It should be emphasised that the existence of harm to the user from cannabis consumption is neither a necessary nor a sufficient condition for policy intervention to be socially desirable. This is an important point, ignored by much of the medical research literature, which concentrates on estimating the magnitude of internal harms, and often appears to assume that the existence of such harms necessarily generates a case for prohibition.

2.1. Direct costs internal to the user

The direct internal costs of cannabis use relate to harms experienced by the user alone. They include damage to health, curtailment of life and loss of productivity. Much depends on the size of these risks, the nature of individual decision-making and the degree to which the risk was appreciated at the time the consumption decisions were made. For economists, the theory of rational addiction (Murphy and Becker 1988) is the benchmark model of demand for a potentially addictive good. It views demand behaviour as fully informed, forward-looking and risk-free and, despite these unpromising assumptions, is able to predict some of the features of drug-dependent behaviour observed among addicts in practice, including price-responsiveness, 'binges' and periods of temporary abstinence. Such behaviour may be a response to adverse external shocks and may be a reflection of great distress, but it is a 'rational' response to those shocks and is, from the addict's point of view, the best possible response. Thus, although policy may have a role through the prevention of events like family dissolution and the creation of employment opportunities, there is no way of improving the welfare of the potential drug user by preventing access to drugs. The Becker-Murphy view has been challenged in recent theoretical work, which suggests a role for policy as a means of improving the welfare of potential drug users. Box 1 indicates the range of behavioural theories that have been proposed and the corresponding policy implications. Note that the latter assume a baseline world without distortions and the implications of these theories in a world with existing tax distortions and market failures is much less clear. The competing theories differ subtly in their conceptualisation of addiction and there remains some doubt about whether cannabis should be regarded as an addictive substance. The work of discriminating between the theories empirically is not well advanced.

Box 1 Behavioural models of addiction and their policy implications

Theory	Features	Internal welfare effects of policy
Rational addiction (Becker and Murphy 1988)	Complementarity between consumption of the addictive good in different periods	No role for policy beyond correction of market failures
Imperfect information (Orphanides and Zervos 1995)	Imperfect information about vulnerability to harm; learning from experience	*Public information* improves welfare, particularly if uncertainty relates to characteristics of the commodity rather than individual vulnerability. Efficient *tax-funded insurance* to fund a drug treatment programmes is welfare-improving. *Prohibition* may increase or decrease welfare.
Temptation (Gul and Pesendorfer 2007)	Preferences relate to both consumption bundles and the opportunity set from which they are chosen, with costly self-control. Consumers maximise the sum of 'commitment' utility and 'temptation' utility.	If temptation is unaffected by non-drug consumption, welfare gains can only be achieved by eliminating temptation, so a pure *tax policy* distorts choice without curtailing temptation and is welfare-decreasing. This result may not hold in more general specifications of temptation. If the commitment utility is maximised with zero drug consumption (i.e. drugs are unambiguously bad), pure *prohibition* (unaccompanied by any price change) improves welfare; otherwise, prohibition may increase or decrease welfare. In general, welfare need not be increased by policies that reduce drug consumption.
Cue-conditioned failures of rationality (Bernheim and Rangel 2005)	Random periods of 'hot' non-rational decision-making, induced by environmental cues; occurrence rates raised by past consumption levels	Policy improves welfare if it *either* reduces probability of 'hot' mode or restrains irrational choices when in hot mode. *Taxes* distort rational choices in 'cold' mode but have no effect on choices in 'hot' mode. *Prohibition* distorts prices causing welfare loss in cold mode decisions and creates supply constraints, with beneficial effects in hot mode but adverse effects in cold mode. *Consumption regulation* improves welfare by reducing environmental cues and varying behaviour in hot mode.
Present-biased decision-making (Gruber and Koszegi 2001)	There are distinct 'selves' in different periods and today's decision may not be optimal for tomorrow's 'self'.	Intra-individual externalities between multiple selves justify *consumption taxes*. If age-specific consumption taxes are impossible, direct regulation of consumption by the young may be justified.
Projection bias (Lowenstein, O'Donoghue and Rabin 2003)	A tendency to under-estimate preference change that will occur in future. Inter-temporal choices are distorted. When craving is strong, future desire is overestimated and quitting is discouraged.	No unambiguous policy implications.

Another more fundamental question is the extent to which all individuals have the cognitive abilities required to make the 'rational' decisions envisaged by intertemporal theories of demand. This is particularly a concern for drug use in adolescence before decision-making skills are fully developed and for people with learning difficulties or other psychological impairments. There is some evidence from survey data of a lack of awareness of drugs and their effects among young teenagers but, paradoxically, ignorance seems to be greatest among those who report no drug use. There is also evidence to suggest that cognitive ability is related to decision-making capacity in general (Burks et al 2008) and to drug use in particular (Conti 2008). However, these findings leave us some way from being able to estimate the share of internal costs which are anticipated and fully considered and therefore not relevant to (non-paternalistic) policy.

How great are the internal potential internal harms in the case of cannabis? Although it is clear that cannabis consumption may have damaging consequences, it is difficult to assemble a strong evidence-based argument that cannabis is any more damaging than several other legal forms of consumption (see MacCoun and Reuter (2001) for a very good overview of the evidence). For example, Nutt et al (2007), who construct an explicit ranking of drugs based on expert opinion, conclude that cannabis should be regarded as significantly less harmful and less addictive than tobacco and alcohol. While there is a large number of people who would be classified (by some measurement conventions) as dependent on cannabis, this is essentially a consequence of the large size of the cannabis user group and there is not much evidence of severe dependency problems (Swift et al 1997).

Virtually no deaths are officially recorded as cannabis-related, but this is partly due to the limitations of the recording process. In many cases it is impossible to tell whether cannabis has contributed to death, since the health impacts of cannabis use are possibly wide-ranging and indirect and not easily identifiable in practice. Possible harms to physical health include lung disease, which is a consequence of the most common form of consumption – smoking, often in combination with tobacco.

The effect of cannabis on the incidence of accidents and violent injury is disputed and some experimental evidence has – rather implausibly – suggested that the calming effect sometimes associated with cannabis may reduce the incidence of some types of accident (Smiley, 1986). Most of the available evidence suggests a positive but quantitatively small association between cannabis use and road accidents (Advisory Council on the Misuse of Drugs 2008), although the evidence is difficult to interpret, not least because many drivers who test positive for cannabis also test positive for alcohol and it is difficult to attribute the excess risk of accident to either substance.

In recent years, most attention has been paid to the link between cannabis consumption and psychotic illness. Despite recent publicity, this is best seen as a continued accumulation of evidence supporting a long-standing concern, rather than a newly-discovered phenomenon or a newly-proved fact (see Arseneault et al 2004 for a review of recent studies). It is certainly true that there is a higher proportion of cannabis users

among people with poor mental health than in the general population and also that cannabis use tends to precede and to have some predictive power for subsequent diagnosis of psychotic and other mental illness (Van Os et al 2002). However, this is a research area beset with confounding factors, making it difficult to determine any causal link between cannabis use and mental illness.

Linked to the mental health concern is recent evidence of change in the production and nature of cannabis products, with a strong rising trend in the average THC-content of cannabis seized by enforcement agencies in the US and Europe: for example a rise from 4.5% to 8.1% in seized samples in the US between 1997 and 2007 (NIDA 2008). This has largely been the result of substitution of domestically-produced sinsemilla for imported lower-potency cannabis. Neither for the US (NIDA 2008) nor the UK (ACMD 2008) is there much evidence of rising potency within each of the hashish/herbal/sinsemilla categories, but the market share of sinsemilla has grown over time. For example, the UK Forensic Science Service reports a rise in the sinsemilla share of seized samples from 15% to 81% over 2002-2008. A large unknown in the link between potency and mental health is the degree to which users adjust their consumption quantities to variations in potency – 'auto-titration', which has long been recognised (CGCINMUD 1972, Mikuriya and Aldrich 1988, Gieringer, 1996). Some volume adjustment is likely (after all, whisky is not generally drunk in the same volume as beer) and it is possible that this effect could reduce some health dangers of cannabis consumption through reductions in the volume of inhaled smoke (Matthias et al. 1997). The lack of any upward trend in hospital admissions for cannabis poisoning matching the trend in potency is consistent with this argument: for example hospital episodes for cannabis poisoning in England fell over the period 2004-7, at a time when average THC potency appears to have been rising rapidly.

The impact of cannabis use on educational attainment, employment and productivity has attracted a great deal of attention from researchers, but with few clear conclusions. Conditional on being in employment, cannabis use has no detectable impact on earnings (and thus, presumably, on productivity), with positive estimated impacts being as common in the research literature as negative ones. Negative associations are generally found between cannabis use (particularly early use) and subsequent educational achievement and employment, but the interpretation of these associations as causal impacts is disputed. For example, Kandel et al (1986) found educational attainment to be unrelated to cannabis use after allowing for observed factors such as pre-existing educational aspirations, whereas Van Ours and Williams (2008), using less informative data but allowing for persistent unobservable confounding factors, found strong effects which they interpret as causal. A common finding casting doubt on causality is that drug use is frequently preceded by other signs of behavioural problems, such as truancy and minor crime (Pudney 2003).

2.2. Direct costs external to the user and policy implementation costs

The few estimates of the external costs of drug use vary widely, ranging from 0.2% of GDP for Canada in 1992 and 0.4% for Germany in 1996 to 1.7% in the US in 2002 and 1.8% for the UK in 2000 (UNDP 1998; ONDCP 2004). These studies differ widely in their coverage and methodology and do not separate cannabis from other drug types. There is no clear distinction between policy implementation costs and other external costs, since many of the latter are a direct consequence of the policy environment – indeed, this is one of the main arguments used by advocates of legalisation. These costs are high: for example, the criminal justice system (CJS) cost of action against drugs in the US is estimated to be over one fifth of total costs (ONDCP 2004). Legalisation of cannabis would remove CJS costs but (to the extent that demand rises in response) , would increase treatment and health care costs. It is difficult to estimate the net effect with any degree of certainty, although the lack of evidence of large adverse health impacts suggests that the net effect of liberalisation will generally be a cost saving.

Drug-related crime constitutes a large part of the external costs of drug use. For example, the UK Drug Harm Index (MacDonald et al., 2005) assigns over two-thirds of its weight to drug-related property crime. For Australia, Collins and Lapsley (2008) estimate that crime was responsible for 56% of the tangible costs arising from illicit drugs in 2004/5; the ONDCP (2004) gives a similar estimate for the US.[2] Drugs might influence crime in various ways. If drug use damages educational achievement and employment opportunity, then acquisitive crime becomes more rewarding relative to legal income-generating activity. Drugs (particularly alcohol) might be directly involved in provoking violent responses to external provocation and violence is also sometimes a feature of the working of illicit markets, where legal enforcement of contracts is impossible. Research on the costs of drug use has generally used crude methods to estimate the proportion of crime regarded as drug-related. One approach is to count as drug-induced all crimes committed by drug users (a counterfactual of perfect law-abidance in the absence of cannabis). It is true that cannabis use is prevalent among offenders: for example, 46% of the 2003/4 England and Wales Arrestee Survey sample reported using cannabis within the previous month (Boreham et al, 2006) compared with 11% for the general population British Crime Survey. However, the difference between these two figures tells us nothing about the causal impact of cannabis use on crime, since there are many potential confounding factors that could produce such an association. A better approach is to use survey questions on motives for criminal activity. For example, using responses to survey questions about criminal motivation, Collins and Lapsley (2008) estimate that 41-51% of crime in Australia in 2004/5 was attributable to drug use. However, this too is open to objection on grounds that criminals themselves may be no better at analysing causation than are social researchers and may find it convenient to blame an external factor for their own behaviour. It is quite likely that many estimates of the incidence and cost of drug-related crime are upward biased for this reason. There

[2] These estimates include policing and CJS costs involved in preventing and responding to drug-related crime.

have been few attempts to estimate causal behavioural models as a basis for simulating crime in a counterfactual drug-free world.

I know of no studies of drug-related crime which separate cannabis from other drug types as a source of crime, partly because many criminals are poly-drug users and it is infeasible to distinguish the separate contribution of each substance. However, there is little reason to believe that the external cost of cannabis-related acquisitive crime is any greater than that of tobacco-related crime, because (like tobacco) cannabis is cheap and available and (unlike tobacco) not believed to be seriously addictive in the way that opiates and cocaine are. Unlike alcohol, there is no credible evidence to support a view of cannabis use as a cause of violent or anti-social behaviour.

The public provision of health care and treatment to drug users who experience personal harm from their consumption is a choice made by a society which perceives the social burden of treatment costs as preferable to leaving drug-induced health problems untreated. There has been a sharp rise in cannabis-related treatment in both the UK and US since the early 1990s (for example, Caulkins et al. (2002) report that caseloads almost doubled in the US between 1993 and 1998) and a similar expansion is under way in the UK. This has been largely a planned expansion in the supply of treatment places, much of which has originated from the criminal justice system, where there is increasing use of schemes trading lighter sentences for enrolment in drug treatment programmes. The cost of drug treatment programmes is: just under a quarter of the US federal budget for drug policy ($3.2bn in 2008) is devoted to treatment and around £0.5bn is budgeted currently by central and local government in England and Wales. It is not possible to allocate these budget totals to individual drugs, but the share of treatment expenditure devoted to cannabis dependency is certainly small.

2.3. Indirect costs – the gateway effect

Given the limited evidence of large direct harms from cannabis consumption, the case for intervention rests principally on indirect harms, internal and external. The possibility of a causal link between cannabis and subsequent hard drug use is known as the gateway or stepping-stone hypothesis and has been the subject of a great deal of research (MacCoun and Reuter 2001, Kandel 2002). The gateway hypothesis holds that the act of consuming cannabis causes an increases in the risk of subsequent use of hard drugs. Note the word “cause” – cannabis and hard drug consumption have a strong statistical association, but the gateway hypothesis goes further than this by asserting that there is a causal relationship responsible for at least part of that association. If a causal gateway does exist, its seriousness depends on the social costs associated with the additional induced hard drug consumption, which are not necessarily large and certainly smaller than the total costs of hard drug use.

The causal gateway effect for a given type of individual can be defined as the difference between two probabilities: the probability of subsequent hard drug use conditional on prior cannabis use and the same probability conditional on no prior

cannabis use. Importantly, the two probabilities should also be conditional on the individual's personal characteristics (observable or unobservable) which are determinants of both types of drug use, so as to isolate cannabis use from these other confounding factors. The problem of causal inference is that, since cannabis use and non-use are mutually exclusive states, only one of the two probabilities can be estimated directly from data – the counterfactual is never observed. There are two ways of overcoming this fundamental problem. One is to use (preferable randomised) controlled experiments: in this case one would administer cannabis to a treatment group and compare their long-run outcomes with those of a control group given no cannabis. Both groups would have to be insulated from the world of cannabis supply and demand but left fully exposed to all other aspects of real life. This is both unethical and impracticable. A second-best approach is observational: use statistical analysis to build a longitudinal behavioural model of the child's development process, including both cannabis and hard drug use as observed outcomes, and then to use the model to simulate lifetime experience of hard drug use in a hypothetical world in which the individual does not experience cannabis.

The major problem here is the impossibility of observing all the personal characteristics that might be common determinants of drug use. With good longitudinal data, we can observe some details of the individual's family history and suitably-designed survey questions can give indicators of cognitive and non-cognitive abilities and impairments. Most surveys fall short of this ideal, but even the most comprehensive longitudinal survey cannot hope to observe every relevant aspect of the individual and his or her environment. Consequently, a major theme of the research literature is the role of common unobserved 'confounding' factors as a source of spurious association between cannabis and hard drug use. In all survey data that I am aware of, it is true that the vast majority (usually well over 90%) of those reporting use of hard drugs (heroin and powder or crack cocaine) also report having used cannabis – almost always earlier than hard drugs. Consequently, naive statistical models of the dynamics of drug use find a strong positive relationship between hard drugs and past cannabis use. But these are not reliable estimates of a causal effect, since precedence in time does not necessarily imply causation – there may be underlying personality traits or family circumstances which are jointly responsible for both forms of drug consumption, with cannabis use occurring earlier because it is more easily available and cheaper than hard drugs.[3] Applied research (e.g. Fergusson and Horwood 2000) has generally found that the inclusion of variables reflecting socio-economic background and psychological attributes reduces the size of the estimated gateway effect, without eliminating it completely.

Many researchers (see Kandel et al 1992, Pudney 2003, Van Ours 2003) have allowed for unobserved confounding factors by assuming them to be fixed over time. It then becomes possible in longitudinal surveys to use the individual's own past to isolate this common unobserved heterogeneity and strip it out, leaving an estimate of the remaining

[3] See Pudney (2003) for a formal analysis

causal impact of cannabis use on hard drug initiation. Studies of this kind also give reduced estimates of the gateway effect, but there often remains a significant effect.

The problem with this approach is that it is not very robust. In particular, the assumption of time-invariant confounding factors is very strong, especially for the group of adolescents and young adults who are the primary concern of drugs research. Recent work on the individual development process emphasises the development through time of abilities and psychological characteristics such as time preference, risk aversion and self-control, rather than their constancy (see Heckman et al 2006).[4] If this is the case, then some component of the common unobserved factors will fail to be removed by constant-effect methods and the element of spurious correlation will appear as an apparently causal gateway effect. Studies of this kind can be expected to overestimate the gateway effect. Box 2 summarises this argument and Appendix A gives a specific technical illustration of the possible bias, using a fixed-effects regression specification as a simple example. The same reservation applies equally to many other areas of drugs research where confounding variables are present, including the impact of drug use on mental health, crime and educational attainment. We return to the gateway hypothesis, particularly its policy implications, in section 3.4 below.

**** BOX 2 HERE ****

3. THE INTERNATIONAL TREATY FRAMEWORK AND DIRECTIONS FOR POLICY CHANGE

3.1. The UN drugs conventions

The system of international drug controls has its origins at the beginning of the twentieth century, with the US as the main driver of the move towards international prohibition of drugs for illicit (meaning non-medicinal and non-research) purposes. The factors contributing to the development of international narcotics law were far from the rational, evidence-based policy design process generally favoured by economists. The first US anti-drug action came in 1905 with Roosevelt's reversal of the plan to restore the old Spanish system of licensed opium supply in the Philippines, following a lobbying campaign by missionaries. Through the 1920s and 1930s, the political and moral forces that led to domestic prohibition of alcohol in the US also drove the US pressure behind international drug control initiatives which led to international acceptance of a drug prohibition objective for both the League of Nations and its successor, the United Nations. A constant thread in this policy development is the emphasis on prohibition of

[4] Note that this developmental view suggests that adverse behavioural outcomes like drug abuse can be seen in part as expressions of developmental problems which may be addressed more effectively by early educational or family interventions than by later drug policy (see Heckman and Materov (2007) for a survey.

non-medical use and criminal penalties for production, trafficking and possession of illicit drugs. This hard line approach remains the basis of international law on drugs.

Cannabis was not originally seen as an object for international drug policy, but was added to the list of proscribed drugs at the League of Nations Second Opium Conference of 1924, following pressure from the Egyptian and Turkish delegations. Cannabis was regarded as a minor issue by many of the participants, who were more concerned with the implementation of controls on the opium trade. Britain and India abstained from the vote, the latter citing cultural and religious reasons for domestic use of the drug. A sub-committee comprising Belgium, Egypt, France, Britain, India, Siam (as was), Turkey and Uruguay produced provisions for the new opium convention, which were adopted unopposed, without significant further consideration. Cannabis has retained its place as an internationally illicit drug since that time, a position not generally reflected in national legislation until rather later.

National governments are currently constrained in their choice of policy on illicit drugs by three international treaties: the 1961 UN Single Convention on Narcotic Drugs (and its 1972 amendment); the 1971 UN Convention on Psychotropic Substances; and the 1988 UN Convention Against Illicit Traffic in Narcotic Drugs and Psychotropic Substances, whose implementation is overseen by the UN International Narcotics Control Board (INCB). The philosophy behind this legal framework was endorsed by the 1998 United Nations General Assembly Special Session (UNGASS), which adopted the hugely implausible slogan “A Drug Free World - We Can Do It!”. Almost 200 countries are signatories to one or more of these conventions, making them among the most successful international agreements in terms of formal adherence. The treaties cover a large number of substances, classified into four groups intended originally to represent their potential for dependence. The schedule I substances subject to complete prohibition include cannabis, cocaine and opiates and synthetics including LSD, amphetamines, barbiturates and MDMA (Ecstasy); schedules II-IV mainly cover medicinal substances subject to less stringent control. There are also measures directed at drug trafficking, money laundering, production of precursor chemicals, and provisions for extradition and asset seizure. These treaties are concerned mainly with the supply of drugs, but Article 3(2) of the 1988 Convention requires nations to make illegal the possession of prohibited drugs for personal consumption under domestic law. Given this international legal framework as a starting point, what are the directions that future policy might take?

3.2. Supply interdiction and consumption policing

Supply interdiction, both domestic and overseas, is the core of current international drugs policy. The dominant player is the US, where the Federal drugs budget devoted to supply-side measures has risen from $5.9bn in 2002 to $8.7bn in 2008 (a rise in the share of the total drugs budget from 55.1% to 63.5%). Of this, $1.08bn rising to $1.67bn has been devoted to overseas interdiction, mostly directed at South American cocaine supply. Many independent commentators regard this as largely ineffective, particularly

in comparison with domestic drug treatment initiatives, which have absorbed a declining share of the Federal budget, despite having a much higher estimated benefit-cost ratio (Rydell and Everingham, 1994). The scope for successful supply-side enforcement policies directed at cannabis production is still more questionable. Unlike cocaine and opiates, production of cannabis is highly dispersed geographically, with the proportion of the product produced domestically in small-scale enterprises believed to be rising in many countries.

Supply-side interdiction can be counter-productive. It is much easier to identify and act against a large organisation - like the Medellin and North Coast cocaine cartels in Colombia, which were centralised enterprises, led by dominant individuals, using violent methods to maintain market dominance. Effective action against them led to the rise of more distributed organisations with a cell-based structure and to still smaller independent producers. Seen from the viewpoint of industrial economics, active supply interdiction tends to have two counteracting effects. First, it increases production costs, by increasing risk and also (since large scale and consequent low unit cost brings visibility) by preventing exploitation of increasing returns to scale. Second, it acts like an anti-trust policy by disrupting large producers and creating market opportunities for smaller competitors, thus tending to increase production and reduce price. It is not immediately clear how these counteracting tendencies play out in practice. Certainly, there is remarkably little evidence of any major long-term effect of supply interdiction in the cocaine market, despite the large scale of resources devoted to it.

On the demand side of the market, there is also little evidence of a large effect of policing. For the typical cannabis user, the risks of apprehension are very low indeed. For example, it has been estimated (see section 4 below) that 5.5 million people used cannabis in England and Wales in 2003/4; approximately 90,000 people[5] were apprehended for cannabis possession, implying an average annual risk of detection around 1.6%. Of course, detection probabilities may be higher for heavy users, who have a greater exposure to risk of detection, but it seems implausible that demand-side policing could have a substantial deterrent effect unless a much stricter and more costly system of enforcement were used.

3.3. Decriminalisation of cannabis use

Terms like legalisation, liberalisation, depenalisation and decriminalisation are widely used in the debate on drugs policy, yet they have no clear definitions and may, as Pacula et al (2004) have pointed out in the US context, obscure large differences between formal and informal enforcement practices. There is a great diversity of forms that relaxation of prohibition can take in practice.

The much-cited case of the Netherlands illustrates the difficulty of classifying enforcement practices under the current international prohibition regime. Cannabis

[5] Author's calculation, from Mwenda and Kumari (2003). The 7% of possession detections where the type of drug was left unspecified have been assumed distributed among drug types in the proportions as those where drug type was recorded.

possession and supply remains formally illegal in the Netherlands but certain outlets are permitted to make small retail sales (up to 5gm.) for consumption on the premises, with strict monitoring to prevent larger sales, underage consumption and sale of harder drugs. In other respects Dutch drug policing is more severe than in many other European countries. In their careful analysis of the Dutch policy, MacCoun and Reuter (2001) identify two phases: an initial post-reform period during which the new freedom had little effect on cannabis demand, followed by a period of increasing commercial exploitation during which the product was more actively promoted and a significant increase in demand may have resulted. However, there is no certainty about this and the scale of cannabis use in the Netherlands is not above the levels observed in the US or several other European countries.

Liberalisation in terms of enforcement of the law on cannabis possession is common internationally. In the US, twelve states are usually identified as decriminalisers of cannabis; Australia, Germany, Italy and Spain have witnessed similar liberalisation. Few analyses of the effects of these policy changes have led to anything but moderate or negligible estimated effects of liberalisation (MacCoun and Reuter 2001, chapter 10).

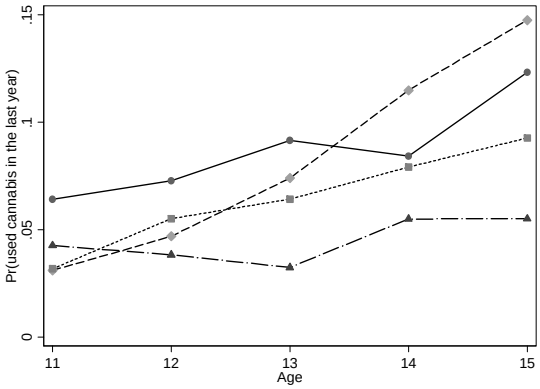
A good example of quasi-decriminalisation is the reclassification of cannabis in the UK. Britain has a 3-category classification system for illegal drugs, with assignment to classes C to A intended to indicate increasing potential for harm. Class C drugs include benzodiazepine and anabolic steroids; class B includes amphetamines and barbiturates, while class A includes opiates and cocaine but also, more contentiously, LSD and Ecstasy. Legal penalties are related to this classification, with maximum jail terms for supply offences being life (class A) and 14 years (classes B and C). Sentences for possession can be up to 7 years (class A), 5 years (class B) and 2 years (class C), although actual sentences are often considerably lighter than these, particularly for own-use possession offences, where on-the-spot seizure and informal warning is the most common outcome for class C drugs.

In January 2004, cannabis was moved from class B to class C status and policing guidelines indicated that, in most instances, possession would not be treated as an arrestable offence. This intervention had three essential features: a possible reduction in the public perception of cannabis harms caused by its move to a lower-harm category; a reduction in the risk involved in the purchase of cannabis (but not in its supply); and a change in policing practices, with a new power for police to seize the drug and give an on-the-spot warning rather than a formal caution following arrest. It was widely expected that these changes might have the effect of increasing cannabis demand and many of us in the research community were waiting eagerly for the availability of survey data from the post-2004 period to analyse the impact of this major policy change. In fact, there is no trace of an upward shift in demand discernible in any of the main UK data sources.

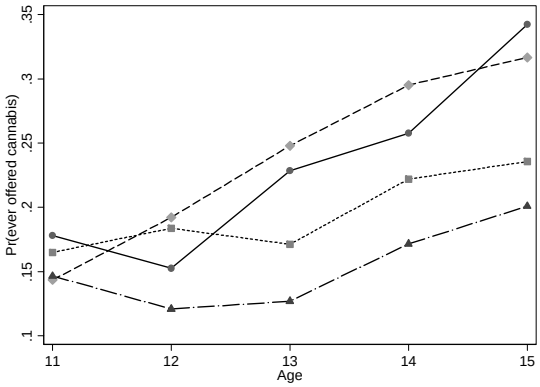
Figure 3 and Table 1 summarise an econometric analysis of the series of annual SSDDU surveys spanning the 2004 policy change. A separate probit model was fitted to each year's data, including dummy variables for age, gender, ethnicity, welfare receipt (entitlement to free school meals) and any episode of exclusion from school. With all other covariates fixed at the pooled sample mean, Figure 3 plots the age profile of the

predicted probabilities of cannabis use, experience of being offered cannabis an acceptability of cannabis, by year. There is a clear fall in these age profiles between the pre-liberalisation period (2001 and 2003) and the post-reform period (2005 and 2007). Any positive impact that reclassification had, either through incentives or informational/social norm signals was evidently small compared to the downward trend in demand, which is most probably the result of a flow of publicity relating to the adverse health effects of cannabis, as reflected in the trend in attitudes (Figure 3c). Paradoxically, by drawing attention to cannabis, liberalisation may have helped to reduce consumption, at least in prevalence terms. Table 1 gives probit coefficients from a single pooled-sample model for each dependent variable, with a test of the hypothesis of no change over time. The downward drift in prevalence, availability and attitudes is highly significant and is also evident in data from the British Crime Survey, which covers a wider age range and is used officially for policy monitoring.

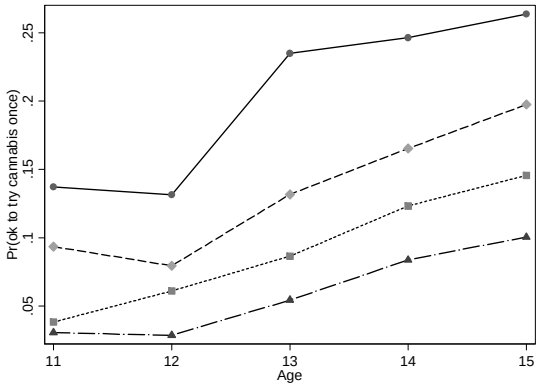
The change in policing practices associated with the British reclassification of cannabis in 2004 is also interesting and perhaps unexpected. Since reclassification there has been a large increase (73% between 2004/5 and 2006/7) in seizures of small quantities of cannabis and a still larger rise (165% between 2004/5 and 2007/8) in the number of police warnings for cannabis possession, which more than offset the corresponding reduction in the number of formal cautions (Kershaw et al 2008). Thus, the probability of *some* penalty – even if only forfeit of the drug and a warning – has risen. One cannot rule out the possibility that this lighter but more extensive pattern of policing contributed to the fall in cannabis prevalence, but the fact that the decline in prevalence began earlier suggests otherwise.



(a) Prevalence: probability of cannabis use in last year



(b) Availability: probability of ever having been offered cannabis



(c) Attitudes: probability of thinking it is "OK to try cannabis just once"

Legend: ——— = 2001; - - - - = 2003; ····· = 2005; -·-·- = 2007

Figure 3. Predicted age profile of prevalence, availability and attitudes to cannabis by year (Year-specific probit models for Survey of Smoking, Drinking and Drug Use among Young People, 2001-2007)

Table 1: Coefficients for probit models of cannabis prevalence, availability and acceptability of cannabis (pooled data, Survey of Smoking, Drinking and Drug Use, 1999-2007; *n* = 53824)

	Used in last year?		Ever been offered?		OK to try once?	
Year 1999	-11.04	***	-4.953	***	-9.062	***
	(1.026)		(0.669)		(1.055)	
Year 2000	-11.00	***	-4.900	***		
	(1.029)		(0.671)			
Year 2001	-10.88	***	-4.922	***	-8.642	***
	(1.025)		(0.670)		(1.053)	
Year 2002	-10.93	***	-4.932	***		
	(1.027)		(0.670)			
Year 2003	-10.92	***	-4.901	***	-8.910	***
	(1.024)		(0.669)		(1.054)	
Year 2004	-11.04	***	-5.002	***	-9.226	***
	(1.027)		(0.671)		(1.055)	
Year 2005	-11.04	***	-5.063	***	-9.168	***
	(1.026)		(0.671)		(1.056)	
Year 2006	-11.19	***	-5.206	***	-9.448	***
	(1.026)		(0.670)		(1.054)	
Year 2007	-11.21	***	-5.215	***	-9.381	***
	(1.027)		(0.670)		(1.055)	
Female	-0.0112		-0.109	***	-0.111	***
	(0.014)		(0.012)		(0.016)	
Mixed race	0.0958	**	0.163	***	0.0906	*
	(0.035)		(0.028)		(0.039)	
Asian	-0.577	***	-0.530	***	-0.491	***
	(0.041)		(0.028)		(0.037)	
Black	-0.282	***	-0.0484		-0.184	***
	(0.048)		(0.037)		(0.055)	
Other	-0.348	***	-0.352	***	-0.248	***
	(0.060)		(0.047)		(0.059)	
Not eligible for free school meals	-0.0655	**	-0.0621	***	0.0176	
	(0.020)		(0.017)		(0.024)	
Never excluded	-0.824	***	-0.760	***	-0.641	***
	(0.020)		(0.018)		(0.022)	
Age	1.134	***	0.310	**	0.839	***
	(0.153)		(0.101)		(0.158)	
Age2	-0.0257	***	0.0056		-0.0145	*
	(0.006)		(0.004)		(0.006)	
$\chi^2(9)$ test of equality of year effects	195.0	***	320.7	***	753.3	***

*; **; *** = significant at 10%, 5% and 1% levels

3.4. Market segmentation

It was argued above that it has not so far been possible to estimate satisfactorily the causal gateway effect linking cannabis use to the risk of hard drugs. Nevertheless, it remains at the heart of the policy debate and one particular form of the gateway hypothesis underlies a class of policies which aim to segment the drugs market and separate cannabis use from other more damaging forms of consumption. This type of policy rests largely on the supply-side or access gateway hypothesis, discussed extensively by Cohen (1972) and adopted as a rationale for the Dutch policy of tacit legalisation. The theory rests on three assumptions: (i) drug users are potentially vulnerable to pressure from others to migrate from cannabis to hard drugs; (ii) a significant number of retail cannabis suppliers are also co-suppliers of harder drugs such as heroin or cocaine; (iii) co-suppliers have a profit incentive to move their clients on from cannabis to harder drugs. Under these conditions, contact with co-suppliers generates a causal gateway effect.

As usual in drugs research, it is difficult to find strong evidence on the existence and magnitude of this effect, but it opens the possibility that, within a prohibition policy regime, the structure of drug supply penalties can have an impact on the rate of migration from cannabis to hard drugs. We illustrate this possibility with some figures from the US and UK. The evidence presented here is necessarily weak and should be seen as illustrative rather than conclusive.

Table 2 shows the structure of formal sentences for trafficking of specific drugs by Federal courts in the US in 2003 and all courts in England and Wales in 2004. It should be emphasised that this is a dangerous comparison, since only a small minority of US cases are tried in Federal courts. However, the differences are suggestive. Jail sentences are more common and on average considerably longer in US Federal courts, partly reflecting the greater seriousness of the average Federal court case. However, there is also a substantial difference in the relative severity of sentences for different drugs. In England and Wales a convicted hard drug supplier can expect a sentence 5.5 times longer than a convicted cannabis supplier. In the US Federal courts, the ratio is only 2.8. If we think of this a relative cost ratio, it suggests that US cannabis suppliers may have a weaker incentive to avoid the hard drug market than their British counterparts. Of course, many other factors will be relevant to these supply decisions, but expected penalties are likely to be important.

This in turn suggests that, if our objective is to reduce the number of cannabis users who migrate to hard drugs, there may be an argument for decreasing, rather than increasing, the penalties for cannabis supply in order to separate the cannabis and hard drug markets and protect cannabis users from pressure from suppliers (a formal analysis of this proposition is given in appendix B). The Dutch ‘coffee shop’ policy can be seen

as an extreme form of this policy, with the cannabis /hard drug penalty ratio set at zero for controlled outlets.

Table 2: Mean sentence lengths for drug offences, US and England & Wales 2004

Drug	US federal drug trafficking convictions, 2003 [1]			England & Wales supply convictions, 2004 [2]		
	Mean nominal sentence (months)	No. imprisoned (thousands)	No. convicted without prison (thousands)	Mean nominal sentence (months)	No. imprisoned (thousands)	No. convicted without prison (thousands)
Cannabis	43.5	4.39	0.826	16.5	1.10	1.98
Cocaine	85.1	3.84		42.6	1.63	0.36
Crack	129.2	3.63		40.4	0.76	0.39
Heroin	66.3	1.23		40.3	2.76	0.66
Mean cannabis sentence [3]	36.6			5.9		
Mean HCC sentence[3]	100.8			32.3		
Ratio HCC / cannabis	2.8			5.5		

[1] Source: USSC (2005), Figure J; figures refer to the period 1 Oct 2003 - 24 Jun 2004. [2] Source: calculated from Home Office (2005), tables S14-S16 and S21; figures refer to calendar 2004. [3] HCC = heroin, crack and cocaine combined. Mean sentences are calculated counting non-custodial sentences as zeros. US figures for the number of zero prison sentences are not broken down by drug type: they have all been assigned to the cannabis category, implying the US HCC/cannabis sentence ratio is an upper bound.

Table 3 continues the US-UK comparison, showing results from the US *Monitoring the Future* survey and the *Youth Lifestyles Survey*, both of which contain questions (listed in Appendix C) on cannabis use and perceived ease of access to hard drugs. The comparison is consistent with the idea that cannabis users in the US have greater exposure to the hard drug market than do cannabis users in the UK. Among non-users, the proportion who report that it would be very easy to obtain hard drugs is 18 % in the US sample but less than half that in the UK. A similar difference is observed for more regular users: for example in the US, 41% of those who use cannabis 40 times or more a year report very easy access , while the figure for similar cannabis users in the UK sample is around 32%. In interpreting this evidence, one should bear in mind the significant design differences between the two surveys and the fact that differences are not necessarily all the result of divergent policies.

Table 3: Ease of access to hard drugs by frequency of cannabis use, US and England & Wales (weighted sample proportions)

Cannabis frequency	Access to hard drugs			
	Very easy	Fairly easy	Fairly hard	Very hard etc.
USA: Monitoring the Future 1990/96/2002 combined				
None	18.3	28.0	23.4	30.2
1-2 occasions	19.5	32.3	26.5	21.8
3-5 occasions	26.4	26.6	24.4	22.5
6-9 occasions	28.7	28.9	24.2	18.2
10-19 occasions	29.2	29.2	23.7	17.9
20-39 occasions	40.4	33.0	11.3	15.4
40 or more	41.0	34.9	14.7	9.4
England & Wales: Youth Lifestyles Survey 1998/9				
None	8.4	14.5	10.4	66.6
Once or twice	14.7	23.4	13.4	48.4
Every couple of months	29.7	17.8	16.1	36.4
Once a month	6.5	32.6	21.7	39.1
2 or 3 times a month	22.1	24.4	14.0	39.5
Once or twice a week	30.9	22.7	13.4	33.0
3-5 days a week	30.4	29.0	17.4	23.2
Every day	40.8	22.5	12.2	24.5

Table 3 estimates the average causal impact of cannabis use on reported ease of access to hard drugs, using four alternative implementations of the propensity score matching method, for two alternative definitions of a cannabis user. In general terms, we take each sampled cannabis user and match him or her with a sampled non-user, chosen to be as similar as possible in terms of a large set of personal characteristics, including measures of access to cannabis, location and neighbourhood characteristics, gender, age, education, ethnicity, smoking and drinking and income.[6] There is a possibility that confounding factors are not fully captured by the set of characteristics used for matching, but, under conditions set out in Appendix C, bias will at least be reduced by matching. The results are consistent with a supply-side gateway effect only in the US sample and only if cannabis users are defined as those who report at least 20 episodes of use in the last 12 months. For that group, the estimated gateway effect is large, with an estimated increase in the probability of easy or very easy access to hard drugs of 31-44 percentage points.

If we choose to believe the supply-side gateway story outlined here, then important policy implications follow. One relates to the design of decriminalisation policies. It is common practice, when possession for personal use is decriminalised, to maintain or

[6] See Appendix C for more detail on the approach and the conditions under which it will reduce the spurious correlation bias inherent in simple comparisons.

even increase the penalties associated with cannabis supply to demonstrate a continued intention to be tough on drugs. In contrast with the Dutch coffee shop policy, the UK reclassification of cannabis in 2004 is an example, when possession became a predominantly non-arrestable offence but supply penalties were retained by raising the maximum supply penalty to 14 years for class C drugs (thus incidentally raising penalties for supply of drugs such as anabolic steroids). This might have been a missed opportunity to segment the market more effectively by reducing the penalties for cannabis supply alongside those for cannabis possession, but the results from Table 3 do not suggest a substantial role for market segmentation in the UK.

Table 3: Propensity score matching estimates of the average effect of cannabis use on the probability of easy/very easy access to hard drugs (t-ratios in parentheses)

	Nearest-neighbour matching	Radius matching	Kernel matching	Stratification matching
USA: Monitoring the Future 1990/96/2002 combined (n = 5,226)				
Any cannabis use in the last year	0.076	0.155	0.043	0.045
	(1.13)	(3.57)	(0.92)	(0.91)
20 or more times a year	0.347	0.438	0.328	0.308
	(3.51)	(7.77)	(5.48)	(4.77)
England & Wales: Youth Lifestyles Survey 1998/9 (n = 3,880)				
Any cannabis use in the last year	-0.048	0.040	-0.044	-0.049
	(1.17)	(1.80)	(1.68)	(1.63)
Cannabis use at least 2 or 3 times a month	-0.007	0.052	0.018	0.007
	(0.15)	(01.78)	(0.58)	(0.21)

3.5. Full legalisation

Prohibition has two effects: on one hand it raises supplier costs, disrupts market functioning and prevents open promotion of the product; on the other, it sacrifices the authorities' ability to tax transactions and regulate operation of the market, product characteristics and promotional activity of suppliers. The cannabis prevalence rates presented in Figure 1 show clearly that prohibition has failed to prevent widespread use of the drug and leaves open the possibility that it might be easier to control the harmful use of cannabis by regulation of a legal market than to control illicit consumption under prohibition. The contrast between the general welcome for tobacco regulation (including bans on smoking in public places) and the deep suspicion of prohibition policy on cannabis is striking and suggests that a middle course of legalised but limited consumption may find a public consensus.

What might a legalised system look like? Excise taxes would almost certainly be at least as high as those on tobacco and alcohol and there would be consequent cross-price effects on those markets. There is mixed research evidence on the direction of these effects, but some studies suggest that cannabis is a complement to tobacco (the two are often consumed in combination) and a substitute for alcohol (see Van Ours 2007), implying that cannabis legalisation would increase the volume of harms from smoking and reduce those from alcohol. However, the magnitude of these impacts is far from clear.

Regulation of a legalised cannabis market could include production and product controls such as bans on publicly accessible production sites and limits on the THC content or (in view of concerns over the possible link with psychotic illness) the balance between THC and CBL content of the marketed product. Product controls of this kind already exist in the tobacco market, where many countries have mandatory testing and disclosure of chemical constituents and Germany and the UK also ban or set quantitative limits on certain tobacco additives (WHO 2001). Product controls could in principle be imposed on legalised cannabis products, but it should be noted that the tobacco industry (perhaps the most likely supplier of legalised cannabis) have a record of resisting regulation very effectively.

Supply regulation would include licensing and inspection of retail sales outlets, with age limits on purchasers. Advertising is likely to be banned or limited, and countered by warning notices and this has been found to have a significant impact on tobacco and alcohol consumption (Saffer and Chaloupka 2000, Saffer and Dave 2002). Consumption in public places would presumably be banned, in line with prevailing policy on tobacco, which has been found to have at least a temporary impact on smoking rates in most countries.

A major advantage of a legalised market is that health information would be separated from political institutions and may consequently become more effective, since there is a much higher level of trust in health workers and advisers than in political institutions and policy-makers. This may be very important: the evidence for the effectiveness of drug prevention programmes under current conditions suggests some limited effect (Caulkins et al 1999), but there is occasional episodic evidence that drug demand is potentially very sensitive to certain kinds of signals about health risk that do not originate from government.[7]

Legalisation would not eliminate the black market completely, since taxation and other controls create opportunities for illicit profit. Illicit trading in tobacco and alcohol is significant in most countries - in the UK, for instance, it is estimated that the market share of smuggled cigarettes has ranged from 20% in 2001/2 to 13% in 2006/7 (HMRC 2008, mid-range estimates). However, the aim of smuggling is generally to evade taxation rather than direct controls on advertising, product quality and consumption, so

[7] The most celebrated example is the death in the US of basketball player Len Bias in 1986 reported by the press to have resulted from his first use of cocaine, followed within a few days by the death of an NFL player, Don Rogers. The media exposure led to an unparalled single-year drop in the proportion of 12th grade students reporting cocaine use and a simultaneous jump in the proportion seeing great risk in using cocaine once or twice (Johnston et al 2006). These effects appear to have persisted for several years.

policy remains partially successful even in the presence of a sizeable black market. Smuggling and the related phenomenon of 'drug tourism' are serious issues when there are large international policy differences, particularly in the EU countries which have freedom of movement and little internal border control.

It is very difficult to predict the outcome of a legalised cannabis market with a system of taxation and regulation in place. It is generally assumed that prices would fall in real terms and consumption rise, although this is not certain. MacCoun and Reuter (2001) have argued in the context of Dutch policy that cannabis use was largely unaffected by the introduction of tacit legalisation in 1976 but may have risen since as a consequence of commercial exploitation by retailers of the market thus created. That is a credible conclusion which suggests that the design of market regulation is a critical issue for the legalisation option. It is likely but not inevitable that legalisation would increase the consumption of cannabis by vulnerable young people but it is not impossible that effective prohibition for the young might prove easier under a legalised system with age controls and effective health information, if commercialisation leads to the curtailment of the illicit supply sector.

4. THE QUALITY OF EVIDENCE – AND A PROPOSAL

Given the difficulty of collecting reliable data on illicit drug use, it is surprising that there is such a large number of surveys which contain questions about drug consumption. For example, since the turn of the century, in the UK (strictly, England and Wales) at least seven different nation-wide surveys[8] have had some coverage of illicit drug use and, in the US, rather more. There are three main problems with the available data resources: scope, non-response and response error. It has been argued convincingly that a good understanding of the demand for drugs requires a broad picture of the child development and socialisation process, with drug use treated as one of the outcomes of that process (Heckman et al, 2006). This implies that we need longitudinal surveys tracking development through childhood into early adulthood, with full observation of the development path. In contrast, most of the large national surveys with a specific drugs focus (like the US *National Survey on Drug Use and Health*, the Australian *National Drug Strategy Household Survey* and the *British Crime Survey*) used for monitoring drug prevalence are conducted as a sequence of cross-sectional surveys (analogous to a set of snapshots rather than a movie). The designers presumably believe that the first priority for research on drug use is a form of prevalence monitoring that only requires detailed questions on drug use, rather than the developmental path leading to drug use - an expensive mistake, in my view.

The problem of poor coverage and non-response stems from the fact that people engaged in illicit activity are less likely to be covered by, and prepared to participate in, a survey involving questions about that activity. This can cause large distortions. For

[8] The *British Crime Survey*, *Arrestee Survey*, *Survey of Smoking, Drinking & Drug Use among Young People*, *British Cohort Survey*, *Longitudinal Study of Young People in England*, *Offending Crime and Justice Survey* and *ONS Survey of Psychiatric Morbidity*. The first three are or were annual cross-sections, the last four are longitudinal.

example, figures from the British Crime Survey (BCS), which interviews people living in private households, suggest that there were around 3.4m cannabis users in England and Wales in 2003/4. A more comprehensive study by Pudney et al. (2006), attempted to overcome the problem of incomplete coverage by the BCS by combining surveys sampling the household population and the flow of police arrestees. The resulting estimates of market size imply a figure of around 5.5m for the number of cannabis users.[9]

The third problem is that survey respondents may give inaccurate answers, with the aim of concealing their drug use from researchers whom they may not trust completely. While it is possible to raise survey response rates and improve response accuracy with interviewing methods (such as computer-assisted self-interviewing) which avoid questioning by an interviewer, these problems cannot be overcome completely and their impact on research findings is not yet well understood.

Even without these data problems, research is not straightforward because of the difficulty of extracting true causal impacts from the statistical associations which are revealed by survey data. Several examples of these inferential difficulties have been identified above and we are still far from a solution.

The major conclusion from a review of the evidence on cannabis markets is – disappointingly – that the evidence is just not good enough to support very strongly any particular view about the best policy to adopt. There certainly are harms from cannabis use but attempts to evaluate them on the basis of individual-level data are frustrated by unobservable confounding factors, misreporting by survey participants and low rates of participation in surveys by groups likely to be particularly vulnerable. The supply side of the market is particularly hard to observe systematically. Experiments with changes in enforcement policy have been made but they are generally local, with cross-contamination between areas, often involve no experimental control at all and their nature is sometimes unclear in practice, partly because the official policy of prohibition must be maintained for conformity with the UN drugs conventions. Because of this last point, policy experiments in practice have also been timid. It has been impossible for any country to take the route of full legalisation with policy focused on taxation and controls on advertising, product quality and location of consumption.

Despite this shortage of good evidence, it is fair to say that, in policy terms, cannabis looks quite different to the other substances covered by the international prohibition treaties. First, it is much more prevalent than drugs like cocaine and heroin and is used routinely (and usually transiently) by a very large number of people who appear to know what they are doing and who suffer no detectable harm. Although there are harms associated with regular use, particularly by very young people, many of this group exhibit other earlier problematic behaviour (Pudney 2003) which suggests the existence of deeper developmental sources of their problems. There is no compelling evidence to suggest that liberalisation would greatly expand this group of potentially vulnerable

[9] The latter figures for numbers of drug users were excluded from the published report by the Home Office, to avoid conflict with the official BCS estimates.

young people and it is even possible to argue that it might be easier to protect this group in a legalised setting.

Perhaps the best way to begin putting policy on a clearer evidential basis is to allow more variation in policy-making, including the legalisation option. This could only be done if domestic supply and consumption of cannabis were to be removed from the international drug prohibition treaties, while retaining the existing ban on international trade in the drug. At the moment we are limited to decriminalisation unaccompanied by the instruments of regulation available for legal markets. This innovation would pass the responsibility for cannabis policy back to national governments, with freedom of action to pursue independent policies. Some will choose to stick to the present prohibitionist framework. Others will choose various forms of decriminalisation and legalisation. We will certainly learn more about the effects of policy.

REFERENCES

Advisory Council on the Misuse of Drugs (2008). Cannabis: classification and public health. London: Home Office.

Arseneault, L., Cannon, M.,Witten J., Murray, R. (2004). Causal association between cannabis and psychosis: examination of the evidence. British Journal of Psychiatry 184, 110–7.

Becker, G. S. and Murphy, K. M. (1988). A theory of rational addiction. Journal of Political Economy, 96, 675-700.

Bernheim, B. D., and A. Rangel (2005). Addiction and Cue-Conditioned Cognitive Processes, American Economic Review 94(5), 1558-1590.

Bohnet, I. and Cooter, R. (2001). Expressive law: framing or equilibrium selection? Berkeley Program in Law and Economics, Working Paper 31.

Boreham, R., Fuller, E., Hills, A. and Pudney, S. E. (2006). The Arrestee Survey Annual Report: Oct 2003 – Sept 2004. England and Wales. London: Home Office Statistical Bulletin 04/06.

Burks, S. V., Carpenter, J. P., Götte, L. and Rustichini, A. (2008). Cognitive Skills Explain Economic Preferences, Strategic Behavior, and Job Attachment. IZA Discussion Paper no. 3609.

Caulkins J., Pacula R., Paddock S., Chiesa J. (2002).School Based Drug Prevention. What Kind of Drug Use Does It Prevent? Santa Monica: RAND Corporation.

Caulkins, J. P., Pacula, R. L., Arkes, J., Reuter, P., Paddock, S., Iguchi, M. and Riley, J. (2004). The Price and Purity of Illicit Drugs: 1981 Through the Second Quarter of 2003. Washington: Office of National Drug Control Policy.

Clements, K. W. (2002). Three Facts about Marijuana Prices. Discussion Paper No. 02.10, Department of Economics, The University of Western Australia.

Cohen, H. (1972). Multiple drug use considered in the light of the stepping-stone hypothesis, International Journal of the Addictions 7, 27-55.

Collins, D. J. and Lapsley, H. M. (2008). The costs of tobacco, alcohol and illicit drug abuse to Australian society in 2004/05. Canberra: Commonwealth of Australia.

Cooter, R. (1998). Expressive law and economics. Journal of Legal Studies 27(s2), 585-607.

Fazey, C. S. J. (2003). The commission on narcotic drugs and the United Nations International Drug Control Programme: politics, policies and prospect for change, International Journal of Drug Policy 14, 155-169.

Fergusson, D. M. and Horwood, L. J. (2000). Does cannabis use encourage other forms of illicit drug use? Addiction, vol. 95, 505-520.

Gul, F. and Pesendorfer, W. (2007). Harmful addiction. Review of Economic Studies 74(1), 147-172.

Gieringer, D. (1996). Marijuana research: waterpipe study, Multidisciplinary Association for Psychedelic Studies Newsletter 6(3) (http://www.maps.org/news-letters/v06n3/06359mj1.html).

Gruber, J. and B. Koszegi (2001). Is Addiction Rational? Theory and Evidence, Quarterly Journal of Economics, 116, 1261-1303.

Gul, F. and Pesendorfer, W. (2007). Harmful addiction, Review of Economic Studies 74(1),147-172.

HM Revenue & Customs (2008). *Tackling Tobacco Smuggling Together*. London: The Stationery Office. (http://www.hmrc.gov.uk/pbr2008/tobacco-2800.pdf)

Hardwick, S. and King, L. (2008). Home Office Cannabis Potency Study 2008. London: Home Office.

Heckman, J. J. and Masterov, D. V. (2007). The productivity argument for investing in young children. IZA Discussion Paper no. 2725

Heckman, J. J., 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.

Johnston, L. D., O'Malley, P. M., Bachman, J. G. and Schulenberg, J. E. (2006). *Monitoring the Future. National survey results on drug use, 1975–2005, Volume II. College Students and Adults Ages 19–45*. Bethesda, MD: National Institute on Drug Abuse.

Kandel, D. B., Yamaguchi, K. and Chen, K. (1992). Stages of progression in drug involvement from adolescence to adulthood: further evidence for the gateway theory, Journal of Studies on Alcohol 53, 447-457.

Kandel, D. B., Davies, M., Karus, D. and Yamaguchi, K. (1986). The consequences in young adulthood of adolescent drug involvement, Archives of General Psychiatry 43, 746-754.

Kandel, D. B. (2002). *Stages and Pathways of Drug Involvement*. Cambridge: Cambridge University Press.

Kershaw, C., Nicholas, S. and Wlaker, A. (2008). *Crime in England and Wales2007/08. Findings from the British Crime Surveyand Police Recorded Crime*. London: Home Office Staistical Bulletin 07/08.

King, L. A., Carpentier, C. and Griffiths, P. (2004). An overview of cannabis potency in European, EMCDDA Insights no. 6. Luxembourg: Office for Official Publications of the European Communities.

Kuran, T. (1998). Ethnic norms and their transformation through reputational cascades. The Journal of Legal Studies 27(s2), 623-659.

Loewenstein, G., O'Donoghue, E. and Rabin, M. (2003). Projection bias in predicting future utility. Quarterly Journal of Economics 118, 1209-1248.

Lynskey, M. T., Heath, A. C., Bucholz, K. K., Slutske, W. S., Madden, P. A. F., Nelson, E. C., Statham, D. J., Martin, N. G. (2003). Escalation of drug use in early-onset cannabis users vs. co-twin controls, Journal of the American Medical Association 298(4), 427-433.

MacCoun, R. J. and Reuter, P. (2001). Drug War Heresies. Learning from Other Vices, Times and Places. Cambridge: Cambridge University Press.

MacDonald, Z., Tinsley, L., Collingwood, J., Jamieson, P., Pudney, S. E. (2005) Measuring the Harm from Illegal Drugs using the Drug Harm Index. London: Home Office Online Report 24/05.

Matthias, P., Tashkin, D. P., Marques-Magallanes, J. A., Wilkins, J. N. and Simmons, M. S. (1997). Effects of varying marijuana potency on deposition of tar and Delta-9 THC in the lung during smoking, Pharmacology, Biochemistry and Behavior 58,1145-1150.

McAllister,W. B. (2000). Drug Diplomacy in the Twentieth Century. London: Routledge.

Mikuriya, T. H. and Aldrich, M. R., (1988). Cannabis 1988. Old drug, new dangers. The potency debate, Journal of Psychoactive Drugs, 20, 47-55.

Mwenda, L. and Kumari, K. (2005). Drug offenders in England and Wales 2003. London, Home Office: Findings 256.

National Institute on Drug Abuse (2008). Potency monitoring project: quarterly report 100, December 16, 2007 thruough March 15, 2008. University of Mississippi.

Nutt, D., King, L. A.,Saulsbury, W., Blakemore, C. (2007). Development of a rational scale to assess the harm of drugs of potential misuse, Lancet 369, 1047-1053.

ONDCP (2004). *The Economic Costs of Drug Abuse in the United States 1992-2002*. Washington DC: Executive Office of the President (publication no. 207303).

Orphanides, A. and Zervos, D. (1995). Rational addiction with learning and regret. Journal of Political Economy 103, 739-758.

Pacula, R., Chriqui, J. F. and King, J. (2004). Marijuana decriminalization: what does it mean in the United States? RAND Working Paper no. WR-126.

Pudney, S. E. (2003). The road to ruin? Sequences of initiation to drugs and crime in Britain, Economic Journal 113, 182-198.

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

Pudney, S.E., Badillo, C., Bryan, M., Burton, J., Conti, G. and Iacovou, M., (2006). Estimating the Size of the UK Illicit Drugs Market, in *Measuring Different Aspects of Problem Drug Use: Methodological Developments*, Home Office Online Report 16/06, pp. 46-120.

Rydell, C. P. and Everingham, S. S. (1994). Controlling Cocaine. Supply Versus Demand Programs.

Saffer, H. and Chaloupka, F. J. (200). The effect of tobacco advertising bans on tobacco consumption. Journal of Health Economics 19, 1117-1137.

Saffer, H. and Dave, D. (2002). Alcohol consumption and alcohol advertising bans. Applied Economics 34, 1325-1334.

Smiley, A. (1986).Marijuana: on-road and driving simulator studies. Alcohol, Drugs and Driving: Abstracts and Reviews 2, 121-134.

Smith, E. S. (1999). The effects of investments in the social capital of youth on political and civic behavior in young adulthood: a longitudinal analysis. Political Psychology 20, 553-580.

Swift, W., Hall, W. and Copeland, J. (1997). Cannabis dependence among long-term users in Sydney, Australia. Australia: National Drug and Alcohol Research Centre.

Torrey-Purta, J., Lehmann, R., Oswald, H. And Schulz, W. (2001). Citizenship and Education in Twenty-eight Countries: Civic Knowledge and Engagement at Age Fourteen. Amsterdam: IEA.

UNDC (1998). *Economic and Social Consequences of Drug Abuse and Illicit Trafficking.* Technical Paper no. 6 (http://www.unodc.org/pdf/technical_series_1998-01-01_1.pdf).

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

Van Ours, J.C. (2007). Cannabis use when it's legal. Addictive Behaviors 32, 1441-1450.

Van Ours, J.C., Williams, J. (2008) Why parents worry: initiation into cannabis use by youth and their educational attainment. Journal of Health Economics, forthcoming.

Van Os, J., Bak, M., Hanssen, M., Bijil, R.V., E Graaf, R., and Verdoux, H. (2002). Cannbais use and psychosis: a longitudinal population based study, American Journal of Epidemiology 156, 319-327.

Warner, J. (2003). Craze. Gin and Debauchery in an Age of Reason, London: Profile Books.

WHO (2001). *Advancing Knowledge on Regulating Tobacco Products*. Geneva: World Health Organisation.

Zuardi, A. W., Shirakawa, I., Finkelfarb, E. and Karniol, I. G. (1982). Action of cannabidiol on the anxiety and other effects produced by Δ^9-THC in normal subjects, Psychopharmacology 76, 245-250.

APPENDIX A. An algebraic example of the fragility of estimates of the gateway effect

Analyses of the gateway effect are usually rather complex, employing methods such as duration analysis (Van Ours 2003), multi-state transition models (Pudney 2003) or other kinds of limited-dependent variable models. These studies allow for unobserved factors specific to the individual and assume them to be time-invariant. The time-invariance assumption is rarely discussed but it is critical to the validity of results from these models. A simpler linear panel regression model will serve to illustrate the issue. Suppose cannabis (*C*) and hard drug (*H*) consumption evolve as follows, with no gateway effect present:

$$C_t = \beta X_{it} + Q_{it} + \varepsilon_{it} \quad \text{(A1)}$$

$$H_t = \gamma X_{it} + \lambda Q_{it} + \eta_{it} \quad \text{(A2)}$$

where X_{it} is a set of observed covariates and ε_{it} and λ_{it} are serially-independent random disturbances, which we assume to be uncorrleated. Q_{it} is a strongly persistent (but not necessarily constant) individual effect. If Q_{it} is time-invariant, then fixed-effects regression offers a good, robust estimator, since it is unaffected by correlation between Q_{it} and X_{it}. Extending the hard drug equation to allow for the possibility of a gateway effect, this involves a regression of $H_{it} - \overline{H}_i$ on $X_{it} - \overline{X}_i$ and $C_{it-1} - \overline{C}_i^{\,l}$, where $\overline{H}_i$ and $\overline{X}_i$ are sample means for individual *i* of observations 1…*T* and the superscript *l* indicates that $\overline{C}_i^{\,l}$ is the mean of the lagged variable, calculated from data in periods 0…*T*-1. If Q_{it} really is time-invariant, this gives a consistent estimate of the true gateway effect of 0.

However, if there is some time-variation in Q_{it}, the lagged consumption variable $C_{it-1} - \overline{C}_i$ contains a random component $Q_{it-1} - \overline{Q}_i^{\,l}$ which is, in general, correlated with the term $\lambda(Q_{it} - \overline{Q}_i)$ which appears in the within-group transformed residual of the hard drugs equation. This correlation is important because the bias in the estimated gateway effect (the coefficient of $Q_{it-1} - \overline{Q}_i^{\,l}$) has the same sign. However, the correlation is complex and depends on the length of the panel, *T*, and the autocorrelation structure of the process $\{Q_{it}\}$. If we assume that Q_{it} follows a 1st-order autoregressive process with parameter ρ, the correlation between $Q_{it-1} - \overline{Q}_i^{\,l}$ and $(Q_{it} - \overline{Q}_i)$ can be derived as a rather messy function of *T* and ρ. To illustrate its implications, Figure A1 plots their correlation against ρ for *T* = 2, 4 and 6.

The conclusion is that, for strong but not perfect persistence of the individual effect (ρ in the neighbourhood of 0.5-0.8), the correlation is positive for *T*>2 and increasing in the length of the panel. Thus, the longer the period spanned by the data, the more serious are the consequences of a departure from time invariance of the individual effect, and the greater the likelihood of finding a spurious positive gateway effect.

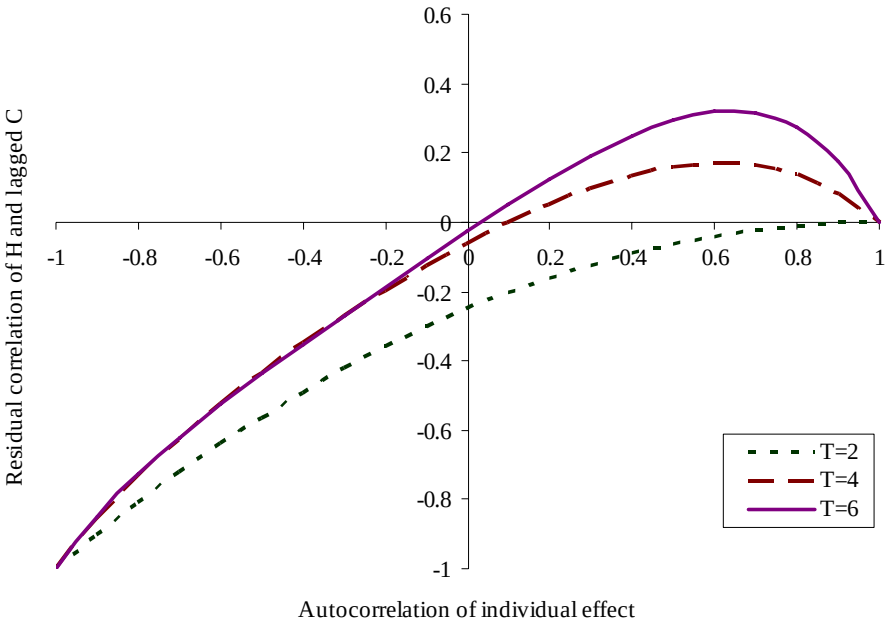


Figure A1. Spurious gateway correlation induced by time-variation in the 'individual effect'. Correlation between $C_{i,t-1} - \overline{C}_i^{\,l}$ and $U_{i,t} - \overline{U}_i$, by autocorrelation of the individual effect (ρ) and panel length (T).

APPENDIX B. The supply-side gateway effect and supply behaviour

Consider a simple fixed-price model in which potential dealers choose between being suppliers of cannabis or hard drugs or both or neither.[10] Dealers face 'standard' penalties $\overline{p}_c$ and $\overline{p}_h$ for cannabis and hard drug supply offences respectively. However, actual penalties are random, since both detection and sentencing are uncertain; $\overline{p}_c$ and $\overline{p}_h$ are thus parameters of the distribution of actual penalties. Potential dealers make their participation decision on the basis of discrete choice stochastic utility maximisation, where U_0,U_c,U_h and U_{ch} are the indirect utilities associated with non-supply, cannabis supply only, hard drug supply only and co-supply of cannabis and hard drugs. Let s be time devoted to legal income-generating activity; c and h are time devoted to cannabis and hard drug dealing, l is non-market time and y is income. Ω_{ch} is the regime-specific budget set:

$$\Omega_{ch} = \{s \geq 0,\ c > 0,\ h > 0,\ l=T-s-c-h,\ y=ws+\pi_c\, c + \pi_h\, h\ \} \qquad \text{(B1)}$$

where T is the time endowment and w is the wage. π_c and π_h are the rates of profit on cannabis and hard drug dealing activity and $\overline{\pi}_c, \overline{\pi}_h$ are the parameters of the distribution of profit rates. Let p be the actual penalty, equal to 0 if unconvicted and some random positive amount if convicted. Consider a co-supplier with ex post utility function $V(y,p)+\varepsilon_{ch}$, where ε_{ch} is an unobservable factor specific to the co-supply regime. Then the supply decision is based on the following indirect expected utility:

$$\max_{\Omega_{ch}} E(V(y,p) \mid s,c,h,l,w,\overline{\pi}_c,\overline{\pi}_h,\overline{p}_c,\overline{p}_h) + \varepsilon_{ch} = \tilde{V}_{ch} + \varepsilon_{ch} \qquad \text{(B2)}$$

Similarly, the other three participation regimes, with suitably defined regime-specific budget sets, yield expected indirect utilities: $\tilde{V}_0 + \varepsilon_0$, $\tilde{V}_c + \varepsilon_c$ and $\tilde{V}_h + \varepsilon_h$, where ε_0, ε_c and ε_h are factors varying randomly across individuals. They could, for example, represent moral attitudes towards soft and hard drugs.

Write the conditional probability that a potential dealer chooses co-supply as Γ. Then $\Gamma = F\left(\tilde{V}_{ch} - \tilde{V}_0, \tilde{V}_{ch} - \tilde{V}_c, \tilde{V}_{ch} - \tilde{V}_h\right)$, where $F(.)$ is the joint distribution function of the three differences ε_{ch} - ε_0, ε_{ch} - ε_c, ε_{ch} - ε_h. Now consider a reform in the Dutch direction, involving a marginal reduction, $d\overline{p}_c$, in the penalty for cannabis supply. Its impact is:

$$d\Gamma = \left((F_1 + F_2 + F_3)\partial\tilde{V}_{ch} / \partial\overline{p}_c - F_2\partial\tilde{V}_c / \partial\overline{p}_c\right)d\overline{p}_c \qquad \text{(B3)}$$

where F_j is the partial derivative of $F(.)$ with respect to its jth argument.

This reform reduces the extent of co-supply if $d\Gamma / d\overline{p}_c > 0$. Since $F_j > 0$, $d\overline{p}_c\partial\tilde{V}_{ch} / \partial\overline{p}_c < 0, \partial\tilde{V}_c / \partial\overline{p}_c < 0$ this requires:

$$\frac{\partial\tilde{V}_{ch} / \partial\overline{p}_c}{\partial\tilde{V}_c / \partial\overline{p}_c} < \frac{F_2}{F_1 + F_2 + F_3} \qquad \text{(B4)}$$

[10] We abstract from the effects of changes in drug prices that might result from supply responses to changing legal penalties.

The left hand side of (B4) is the ratio of the marginal expected (dis)utilities of the cannabis penalty in the co-supply and cannabis-only regimes, while the right hand side is the marginal response of the co-supply probability with respect to $\tilde{V}_c$ as a share of the responses to $\tilde{V}_0, \tilde{V}_c$ and $\tilde{V}_h$. Consequently, for a given individual, the likelihood of co-supply can be reduced by a cut in the penalty for cannabis supply provided cannabis supply is a significant alternative to co-supply (in the sense that $F_2 / (F_1+F_2+F_3)$ is large) and that marginal disutility of the penalty is substantially higher in the cannabis-only than the co-supply mode. A necessary condition for (B4) to be satisfied is that $|\partial\tilde{V}_{ch} / \partial\bar{p}_c| > |\partial\tilde{V}_c / \partial\bar{p}_c|$. This is not automatically implied by concavity. For example, consider the simple case where penalties have a cash-equivalent, so that $V(y, p)=V(y-p)$. Assume there are regime-specific income levels y_c and y_{ch} (with $y_{ch} > y_c$), known with certainty, a constant conviction probability θ and, following conviction, known penalty levels of p_c and p_{ch} (where $p_{ch} > p_c$). Then $|\partial\tilde{V}_{ch} / \partial\bar{p}_c| > |\partial\tilde{V}_c / \partial\bar{p}_c|$ if and only if :

$$(1-\theta)\left(V'(y_{ch}) - V'(y_c)\right) + \theta\left(V'(y_{ch} - p_{ch}) - V'(y_c - p_c)\right) < 0 \qquad \text{(B5)}$$

This will be satisfied if $V(.)$ is concave and $y_{ch} - p_{ch} > y_c - p_c$, but also in cases where $y_{ch} - p_{ch} < y_c - p_c$, provided the conviction probability, θ, is sufficiently small.

Note that the number of hard drug dealers will rise since $F^*\left(\tilde{V}_h - \tilde{V}_0, \tilde{V}_h - \tilde{V}_c, \tilde{V}_h - \tilde{V}_{ch}\right)$ is increasing in $\bar{p}_c$, where F^* is the distribution function of ($\varepsilon_h - \varepsilon_0$, $\varepsilon_h - \varepsilon_c$, $\varepsilon_h - \varepsilon_{ch}$), so increased competition in the hard drug market may lead to some increase in demand by hard drug users, induced by price falls.

Now assume that consumers of cannabis are vulnerable to social influence by their suppliers, who have a financial incentive to sell hard drugs in preference to cannabis (since $\pi_h > \pi_c$). Each cannabis user locates a source of cannabis supply through a random matching process in which suppliers are drawn at random from the population of dealers offering cannabis. If the dealer turns out to be a co-supplier, there is then a probability Q that the consumer will subsequently succumb to pressure and make the transition to hard drugs; otherwise, the consumer remains a cannabis user only. Let the probability of being a cannabis user be a decreasing function, $\varphi(R)$ of the penalty R for cannabis possession. Then the number of cannabis users who move to hard drugs is proportional to $\varphi(R)Q$ E(Γ), where the expectation is with respect to the distribution of w, π_h, π_c in the population of potential dealers. The effect on the prevalence of hard drug use of simultaneously varying the penalties for cannabis possession (dR) and cannabis supply ($d\bar{p}_c$) is:

$$\text{d Pr (hard drug use)} = Q\left\{\varphi'(R)E(\Gamma)dR - \varphi(R)[\partial E(\Gamma) / \partial\bar{p}_c]d\bar{p}_c\right\} \qquad \text{(B6)}$$

Hard drug prevalence will decline if:

$$\text{d ln(R)} > \left[\frac{\partial \ln E(\Gamma) / \partial\bar{p}_c}{\partial \ln\varphi(R) / \partial R}\right] d\ln\bar{p}_c \qquad \text{(B7)}$$

A reform in the direction of the British 2004 reclassification, leaving unchanged the penalty for cannabis supply ($d\bar{p}_c = 0$) but reducing the penalty for possession ($d \ln R < 0$) will unambiguously increase hard drug prevalence under these assumptions. A reform in the Dutch direction reduces both penalties ($d \ln \bar{p}_c, d \ln R < 0$) and can therefore reduce prevalence provided the term in square brackets in (B7) is positive. This in turn requires that $\partial \ln E(\Gamma) / \partial \bar{p}_c < 0$, which is satisfied if the condition for co-supply (B4) characterises a sufficiently large part of the potential supplier population.

APPENDIX C. The matching method

Define measures of cannabis consumption, *C*, and hard drug access, A_h. These are constructed as binary variables respectively indicating: consumption of cannabis above a specified threshold frequency; and easy or very easy access to any hard drug. Define a pair of latent outcomes for access to hard drugs, (A_h^1, A_h^0) which respectively would be realised in the circumstances of cannabis use ($C = 1$) or abstention ($C = 0$). Only a single outcome is actually observed: $A_h = A_h^1$ if $C = 1$ and A_h^0 if $C = 1)$. For cannabis users, the impact of their cannabis consumption on their exposure to hard drugs is:

$$\Delta = E\left(A_h^1 - A_h^0 \mid C = 1\right) \tag{C1}$$

In the terminology of programme evaluation, Δis the average effect of treatment on the treated, where we interpret treatment to mean cannabis use. Follow Rosenbaum and Rubin (1983) in assuming (A_h^1, A_h^0) are independent of cannabis use *C*, after conditioning on a suitable set of observable covariates, *X*. Under these conditions, the expected value in (C1) is:

$$\Delta = E\left(E\left(A_h^1 \mid C = 1, p(X)\right) - E\left(A_h^0 \mid C = 0, p(X)\right) \mid C = 0\right) \tag{C2}$$

where $pI(X) = \Pr(C=1|X)$ is the propensity score, defined as the conditional probability of cannabis use.

The choice of variables, *X*, to include in the selection model is important. The matching estimator only controls for selection on observables, so it is important to use covariates that capture as far as possible any unobservable factors which connect the decision to become a cannabis user with the availability of hard drugs. In this sense, the approach to modelling selection is unusual, since it should aim to use variables 'endogenous' to selection. The selection model should not be thought of a structural behavioural model, where we would normally seek to avoid endogeneity bias. Assume that observable variables *Z* and unobservables *U* influence both cannabis use, *C*, and hard drug availability, (A_h^1, A_h^0). Our objective is to determine whether realised access A_h is also influenced directly by *C*. Assume that a further set of variables, *W*, is also determined by *Z* and *U* and is, in that sense, endogenous. However *W* is known a priori not to be directly influenced by *C*.

The formal assumptions corresponding to this causal structure are:

$$A_h^0 \perp C \mid Z, U\,; \qquad W \perp (A_h^0, C) \mid Z, U \tag{C3}$$

where ⊥ denotes statistical independence. If we use $X=(Z, W)$ as covariates in the matching analysis, then there will be no bias if we can show that $(A_h^0, A_h^1) \perp C \mid Z, W$. For simplicity, assume that *U* is discrete and consider the distribution $P(C, A_h^0, A_h^1 \mid Z, W)$, which can be written:

$$P(C, A_h^0, A_h^1 \mid Z, W) = P(C, A_h^0, A_h^1 \mid Z, W) / P(W \mid Z)$$
$$= \frac{\sum_u P(C \mid Z, u) P(A_h^0, A_h^1 \mid Z, u) P(W \mid Z, u) P(u \mid Z)}{\sum_u P(W \mid Z, u) P(u \mid Z)} \quad \text{(C4)}$$

The conditional independence property $(A_h^0, A_h^1) \perp C \mid Z, W$ is then satisfied if (\ref{decomposition}) can be written as the product of two components: one a function of *C, Z, W*, the other a function of A_h^0, A_h^1, Z, W. Apart from the trivial case where *C* or A_h^0, A_h^1 are independent of *U*, there are two alternative assumptions that will generate independence conditional on *W, Z*.

Assumption 1 At all points in the support of *W, Z*, there exists a unique function $\tilde{u}(Z, W)$ such that $P(W \mid Z, U) = 0 \text{ for all } U \neq \tilde{u}(Z, W)$.

Assumption 1 essentially states that, given *Z, W* is an exact, invertible function of *U*. In this case there is only a single non-zero term in the sums in numerator and denominator of (C3), so $P(C, A_h^0, A_h^1 \mid Z, W)$ reduces to $P(C \mid Z, \tilde{u}(W, Z)) P(A_h^0, A_h^1 \mid Z, \tilde{u}(W, Z))$. The alternative assumption is more complex. Define a variable:

$$\lambda(u \mid W, Z) = \frac{P(W \mid Z, u) P(u \mid Z)}{\sum_{u^*} P(W \mid Z, u^*) P(u^* \mid Z)} \quad \text{(C5)}$$

Assumption 2 At all points in the support of *W, Z*, there exists a pair of functions $\lambda_c(u|W, Z)$, $\lambda_a(u|W, Z)$ such that $\lambda_c = \lambda_c \lambda_a$ and:

$$\sum_u \left(P(C \mid Z, u) \lambda_c(u \mid W, Z) - \mu_c\right)\left(P(A_h^0, A_h^1 \mid Z, u) \lambda_a(u \mid W, Z) - \mu_a\right) = 0 \quad \text{(C6)}$$

where $\mu_c(C, W, Z) = \sum_u P(C \mid Z, u) \lambda_c(u \mid W, Z)$ and $\mu_a(A_h^0, A_h^1, W, Z) = \sum_u P(A_h^0, A_h^1 \mid Z, u) \lambda_c(u \mid W, Z)$

Under assumption 2, $P(C, A_h^0, A_h^1 \mid Z, W) = \mu_c(C, W, Z) \mu_a(A_h^0, A_h^1, Z, W)$, which implies conditional independence of $C, \text{and } (A_h^0, A_h^1)$. Note the role of the variables *W* here. If *U* influences $C \text{ and } (A_h^0, A_h^1)$ positively (say), then without the terms λ_c and λ_a assumption (C5) would be violated. However, use of the variables *W* makes it possible to introduce negatively-covarying λ_c and λ_a which offset the positive covariation between $P(C \mid Z, u)$ and $P(A_h^0, A_h^1 \mid Z, u)$. Thus, suitable 'endogenous' covariates can be effective in reducing the extent of selection bias.

In our implementation of the matching estimator, *X* includes variables describing behaviour closely related to cannabis use (including smoking, drunkenness and various illicit behaviours) and also access to cannabis, A_c. These are almost certainly endogenous, in the sense that they are influenced by the same unobservable personal factors that affect cannabis consumption and accessibility of hard drugs. However, it is plausible to assume there is no direct impact on them of cannabis use. Drinking, smoking and low-level illicit behaviour generally pre-date cannabis consumption (see Pudney, 2003) and accessibility of cannabis is logically prior to its consumption.

Provided we can match each cannabis user to a non-user with a similar value of the propensity score, the contribution of cannabis use to hard drug exposure is estimated as the difference between the means of the observed variable A_h for the treatments and controls. Thus, in the simplest case of one-to-one matching, we can estimate (C1) as:

$$\hat{\Delta} = \overline{A}_h^1 - \overline{A}_h^0 \qquad \text{(C6)}$$

where $\overline{A}_h^1$ and $\overline{A}_h^0$ are the sample means of A_h among the sampled cannabis users and the matched non-users respectively.

We use the Becker and Ichino (2002) implementation of the matching estimator. The first stage of this process involves estimating the propensity score using a probit model. This is then tested, using the region of common support of the treatment and control cases, to check that the balancing property $C \perp X \mid p(X)$ is satisfied. One version of the analysis defines C as an indicator of whether there has been any cannabis use in the last year; in the other, C indicates whether there has been at least 20 episodes of use in the last year (MtF) or use at least 2 or 3 times a month in the last year (YLS). The probit models used to generate the propensity score are set out in Table C2 and C3; all satisfy the balancing property.

An important issue is the degree of overlap between the treatment and control groups. To ensure at least one match for all treatment cases, we need the condition $N_t\, f_t(p) \leq N_c\, f_c(p)$ where N_t and N_c are the numbers of treatment and control cases and $f_t(p)$ and $f_c(p)$ are the densities of the propensity score in the two groups. This condition is generally satisfied, except for the region of large (above 0.5 for the MtF or 0.75 for YLS) propensity scores for the 'any cannabis use in the last year' definition of C. The implication of this exception to the common support condition is that members of the control group with large propensity scores are used to match more than one treatment case. There is only a small region where $f_t(p) > f_c(p) = 0$ (implying that there are no matches at all). Thus, when we impose the common support restriction, which discards treatment cases with no acceptable match, we lose a negligible number of cases from the analysis and the results are not perceptibly affected.

The MtF and YLS Drug Availability Questions

The principal survey questions about cannabis use and access to drugs in the MtF and YLS (identified by their codes in the survey documentation) are as follows.

MtF questions

V2116 "On how many occasions (if any) have you used marijuana (weed, pot) or hashish (hash, hash oil) during the last 12 months?"

The responses are coded as: 1 "0 occasions"; 2 "1 or 2 occasions"; 3 "3-5 occasions; 4 "6-9 occasions"; 5 "10-19 occasions"; 6 "20-39 occasions"; 7 "40 or more"

V2464 "How difficult do you think it would be for you to get each of the following, if you wanted some? A: "Crack" cocaine

V2465 "How difficult do you think it would be for you to get each of the following, if you wanted some B: Cocaine in powder form

V2309 "How difficult do you think it would be for you to get each of the following, if you wanted some H: Heroin

The responses to V2464, V2465 and V2309 are separately coded as: 5 "very easy"; 4 "fairly easy"; 3 "fairly difficult"; 2 "very difficult"; and 1 "probably impossible". We then construct a variable equal to the maximum of these scores over the three questions.

YLS questions

CFREQ2 "How often have you taken Cannabis (Marijuana, Grass, Hash, Ganja, Blow, Draw, Skunk) in the last 12 months?"

The responses are coded as: 0 Not at all; 1 Once or twice this year; 2 Once every couple of months; 3 Once a month; 4 Two or three times a month; 5 Once or twice a week; 6 Three to five days a week; 7 Every day.

CACCESS2 "If you wanted to get cannabis (...) and you had the time and money to do so, how easy would it be?"

The responses are coded as: 1 don't know or very difficult; 2 fairly difficult; 3 fairly easy; 4 very easy.

CACCESS6 "And how easy would it be for you to get heroin (smack, skag, `H', brown)?"

CACCESS3 "And how easy would it be for you to get cocaine (coke)?"

The responses to these last two variables are coded into a single variable on the basis of whichever drug has the easier availability: 4 very easy; 3 fairly easy; 2 fairly hard; 1 very hard or don't know.

Table C1: MtF and YLS weighted sample characteristics

MtF data (n = 5226)		**YLS data** (n = 3880)	
Variable	Sample mean / proportion	Variable	Sample mean / proportion
North East	0.206	North	0.048
North Central	0.298	Yorkshire & Humberside	0.075
South	0.360	North West	0.086
Small-medium town	0.476	East Midlands	0.063
Large city	0.159	West Midlands	0.079
Aged 18 or over	0.664	East Anglia	0.033
B grade school work	0.505	South East	0.156
A grade school work	0.280	South West	0.060
Female	0.497	Wales	0.044
Black	0.095	Rural	0.200
Father high school	0.480	Inner city	0.147
Father college	0.362	Deprived area	0.066
Father education not recorded/not applicable	0.028	Age	21.1
Absent parent	0.263	Born in UK	0.943
Smoker	0.230	Left school at or before 16	0.310
Drinker	0.461	University/college student	0.058
Income	74.64	Still in full-time education	0.391
Zero income	0.053	Years since education completed [1]	4.39
1990	0.391	Ever expelled from school	0.017
1996	0.360	Ever in care	0.022
		Female	0.501
		Black	0.019
		Asian	0.043
		Absent parent	0.074
		Father unemployed	0.021
		Deprived neighbourhood	0.066
		Family trouble with police	0.018
		Currently smokes	0.306
		Ever been drunk	0.551
		Dislikes school	0.096
		Hangs around town	0.111
		Little parental supervision	0.256

[1] Conditional mean for respondents who have completed full-time education

Table C2 Propensity score probit model for US MtF data

Covariate	Any cannabis use in last year		Use at least 2 or 3 times a month	
	Coefficient	**Std. Err.**	**Coefficient**	**Std. Err.**
Very easy cannabis access	-1.869	0.288	-1.239	0.452
Fairly easy cannabis access	-1.170	0.238	-1.038	0.275
Fairly difficult cannabis access	-1.025	0.135		
Very difficult cannabis access	-0.500	0.052	-0.546	0.087
North-East	-0.142	0.072	-0.054	0.097
North Central	-0.107	0.071	-0.086	0.098
South	-0.239	0.070	-0.148	0.098
Large city	0.056	0.059	0.052	0.082
Over 18	-0.050	0.046	0.041	0.062
B-grade school work	-0.180	0.054	-0.196	0.069
A-grade school work	-0.261	0.065	-0.321	0.089
Female	-0.004	0.044	-0.244	0.061
Black	-0.021	0.079	-0.104	0.125
Father high school	-0.064	0.068	-0.145	0.091
Father college	-0.008	0.071	-0.113	0.095
Father's education not recorded	-0.078	0.142	0.042	0.184
Absent parent	0.117	0.049	0.070	0.065
Smoker	0.823	0.050	0.843	0.060
Drinker	1.126	0.045	1.034	0.078
Income	0.001	0.000	0.002	0.000
Zero income	-0.211	0.118	0.139	0.168
1990	-0.280	0.058	-0.252	0.080
1996	-0.090	0.057	0.054	0.076
Intercept	-0.674	0.119	-1.878	0.169
Common support	(0.00664, 0.94324)		(0.00228, 0.73393)	

Note: Variables listed in Table C1 which do not appear above were insignificant at the 10% level and were dropped from the model.

Table C3 Propensity score probit model for Engalnd & Wales YLS data

Covariate	Any cannabis use in last year		Use at least 2 or 3 times a month	
	Coefficient	**Std. Err.**	**Coefficient**	**Std. Err.**
Very easy cannabis access	1.432	0.094	1.868	0.242
Fairly easy cannabis access	0.904	0.102	1.218	0.251
Fairly difficult cannabis access	0.582	0.171	1.116	0.319
10/age	-3.998	0.497	11.613	4.205
$(10/age)^2$	-	-	-13.482	3.946
Still in full-time education	-0.232	0.101	-0.124	0.125
Left school at or before age 16	-0.073	0.079	-	-
Female	-0.254	0.059	-0.382	0.077
Asian	-0.549	0.244	-	-
Absent parent	0.233	0.099	0.235	0.122
Deprived neighbourhood	0.186	0.096	0.225	0.116
Smoker × (10/age)	2.091	0.127	2.032	0.165
Has been drunk × (10/age)	1.380	0.146	0.630	0.203
Dislikes school	0.123	0.087	0.183	0.102
Hangs around town	0.476	0.111	0.489	0.134
Little parental supervision	0.169	0.063	0.265	0.077
Years since leaving education	-0.092	0.012	-0.038	0.017
Intercept	-0.103	0.272	-5.572	1.165
Common support	(0.00071, 0.96256)		(0.00110, 0.87722)	

Note: Variables listed in Table C1 which do not appear above were insignificant at the 10% level and were dropped from the model.