

HEDG Working Paper 05/09

Sequential patterns of drug use initiation – can
we believe in the gateway theory?

Anne Line Bretteville-Jensen
Hans Olav Melberg
Andrew M Jones

October 2005

ISSN 1751-1976

Sequential patterns of drug use initiation – can we believe in the gateway theory?

Anne Line Bretteville-Jensen*

Hans Olav Melberg**

Andrew M. Jones***

Abstract

The gateway, or stepping stone, hypothesis is important as it has had considerable influence on drug policy and legislation in many countries. The gateway hypothesis offers one possible explanation for young people's development of a serious drug problem. It simply states that the use of one drug increases the risk of starting to consume another, and possibly more harmful, drug later on and that the risk increases with frequency of use (dose-response). The empirical basis for the hypothesis is the common finding that most heavy drug users have started with less dangerous drugs first and that there seems to be a "staircase" from alcohol and insolvents via cannabis and tablets to amphetamine, cocaine and heroin. The core question is whether the sequential initiation pattern of drug use is best explained by the mechanisms substantiating the gateway hypothesis or whether the phenomenon is better understood by employing the concepts of accessibility and/or transition proneness? Based on a representative sample of 21-31 year olds in Oslo we have examined the possible gateway effect of both legal (alcohol) and illegal drugs (cannabis) on subsequent use of cannabis and hard drugs (amphetamine and cocaine). We use multivariate probit models that take account of unobservable individual-specific effects to reduce the possibility of a spurious causal effect of soft drug use on the onset of hard drug use. The gateway effects were greater when we did not take account of unobserved heterogeneity, but, although substantially reduced, they remained considerable also when unobserved factors were accounted for.

JEL: I12, I18

Keywords: Gateway hypothesis, Stepping stone hypothesis, Substance abuse, Multivariate probit analysis, Unobserved heterogeneity, Amphetamine, Cocaine, Cannabis

*Norwegian Institute for Alcohol and Drug Research (SIRUS)
and Health Economics Bergen (HEB)

** Norwegian Institute for Alcohol and Drug Research (SIRUS)

***Department of Economics and Related Studies, University of York
and Health Economics Bergen (HEB)

1 Introduction

The gateway hypothesis offers one possible explanation for young people's development of a serious drug problem. It simply states that the use of one drug increases the risk of starting to consume another, and possibly more harmful, drug later on and that the risk increases with frequency of use (dose-response). The empirical basis for the hypothesis is the common finding that most heavy drug users have started with less dangerous drugs first and that there seems to be a "staircase" from alcohol and solvents via cannabis and tablets to amphetamine, cocaine and heroin. With cannabis being the first illegal drug for most drug users it has been denoted "the gateway drug".

The gateway, or stepping stone, hypothesis is important as it has had considerable influence on drug policy and legislation in many countries. The strict penal sanctions directed against cannabis use may at least in parts be explained by the fear for users' possible progression to harder drugs, and the gateway effect is often used as an argument against cannabis legalisation. A strong belief in the gateway theory is revealed in many studies, for example:

“And marijuana, I contend, was the drug that brought many young people across that psychological boundary of doing something that was illegal and illicit, based on predominant norms. It was the pathbreaking drug that tore a great hole through the fabric of traditional normative social constraint and made it far easier for young people to consider using other drugs.” (Johnston, 1991, p. 107)

Not every country, however, has adopted the strict approach and some have recently changed theirs in a more liberal direction. In contrast to most other Western societies, the Netherlands has for long aimed at separating the markets for soft and hard drugs by allowing "coffee shops" to legally sell cannabis while keeping a strict regime against the trade of hard drugs like cocaine and heroin. The Netherlands has been criticised for their stance but more recently other countries like Portugal and the UK have changed their drug legislation too. Re-classifying cannabis to a class-C drug, as was implemented in the UK in 2004, is one step in the direction of de-criminalizing the drug. Whether this change subsequently will lead to an increase in the number of heavy drug users depends, among other things, on whether there is a gateway effect or not. Is the sequential pattern commonly observed only mirroring the path to heavy drug use that is influenced by accessibility (physical, legal, economic, cultural) without there being any casual link between the different stages?

The present paper aims at examining the gateway effect by analysing data collected among the general population of 21-31 year olds in Oslo. We employ models that take account of unobservable individual-specific effects to reduce the possibility of a spurious causal effect of soft drug use on the onset of hard drug use. Many youngsters seem to experiment with illicit drugs. The majority of problems related to drug use, however, are caused by regular users. Therefore, and in contrast to most studies in this field, we examine the gateway hypothesis studying “users” and separate people according to their frequency of use, not according to whether or not they report to have ever tried various drugs. The most policy-relevant question is not whether a soft drug makes it more likely that a person will just try a hard drug once at some later point in time, but whether having used a soft drug makes it more likely that the individual will progress to have a problematic use of a hard drugs later. Before presenting the methods, data and results, however, we look more into the gateway theory and give an overview of the relevant empirical literature.

2 The gateway theory

Kandel (1975) was one of the first to study the sequential pattern of drug use initiation based on longitudinal data. She found four stages in drug use with marijuana being the crucial step on the path to other illicit drugs. As is emphasised in Kandel *et al.* (1992), however, the authors state that entry into a particular stage is common and perhaps even necessary although not a sufficient prerequisite for entry into the next stage, i.e. they argue against a version of the stepping stone theory which claims that marijuana leads inexorably to the use of other illicit drugs. Goode (1972), on the other hand, is an early example of the assertion that there could be a casual link between different stages in drug use.

MacCoun and Reuter (2001) have a thorough discussion of the concept and Pudney (2003) lists three possible mechanisms that, on their own or in combination, might be the basis for a causal gateway effect in drug use:

1. The consumption of soft drugs may create a psychological or physiological need for further and stronger experiences of the same type.
2. The act of obtaining and using soft drugs may bring the user into contact with hard-drug users or suppliers whom they would not otherwise have met.

3. Experience of the use of soft drugs with no obvious ill effects may appear to contradict and undermine the strong negative publicity directed against illicit drug use in general, so that advice against hard drugs becomes less persuasive.

In addition to these mechanisms, one could argue that for some individuals consuming a drug for the first time is like crossing a threshold and that the action makes it less costly to proceed into another "drug stage". Taking only one step at a time, when each step reduces the cost of the next, could increase the probability of ending up as a heavy drug user. Some people may not have started to consume, for instance heroin, if they were offered the drug without first having tried other illegal substances. Hence, despite increased legal sanctions and increased dangers associated with consumption of the various drugs along the path, some people proceed, and the claim is that there is a causal effect through reduced costs caused by the initial consumption of drugs at each stage.

Pacula (1997) approaches the possible causal relationship from a different angle. In contrast to reduced costs, she suggests that past consumption of any one drug will increase the marginal utility of consuming this drug and any other drug. Her model is a variation of the rational addiction framework developed by Becker and Murphy (1988) and it builds in reinforcement and tolerance effects of addictive goods. Pacula differs from Becker and Murphy in that she assumes that the "consumption capital" represents past consumption of many substances, and by that she opens the way for a possible gateway effect of drugs. She claims that young people start consuming the drug with the lowest marginal cost and then become more likely to initiate use of more costly substances as the marginal utility of using them rises. In Kenkel *et al.* (2001) the authors describe the data that would be required for a proper test of the rational addiction version of the gateway hypothesis.

If the use of soft drugs has a causal gateway effect on the transition to harder drugs, then restrictions on the use of soft drugs may be an effective policy tool to achieve the objective of reducing the use of hard drugs. However there are alternative explanations for the observed data that do not have the same policy implications. An alternative to the gateway hypothesis for explaining the observed sequential pattern of drug initiation is *differences in accessibility* for various age groups. With accessibility we mean physical availability, cultural acceptance, prevailing drug legislation and affordability (influenced by both individual income and drug prices). People may start to consume alcohol prior to cannabis and cannabis

prior to cocaine simply because the former is more accessible to very young people. This is in line with general economic theory. It seems obvious that the varying accessibility of different drugs to young people at different stages in their adolescence influence the observed drug using pattern and, as such, no causal effect is needed to explain the path to hard drug use.

Another alternative is "transition proneness" (see e.g. Jessor *et al.* (1980) for an early version of this hypothesis). The claim is that there is a pattern of proneness or vulnerability to deviance among people consuming illegal drugs. Using drugs is only one response to certain environmental conditions or personal characteristics which may result in a wide range of deviant behaviour. Individuals may differ greatly with respect to such an influence, and empirically, relevant information on proneness is hard to reveal through general surveys or in-depth interviews. Thus, the observed correlations may be spurious, reflecting some third factor that influences the use of several types of drugs. The literature refers to un-revealed information among observation units as unobserved heterogeneity.

Morral *et al.* (2002) construct a model to test whether the gateway effect could be explained by such a third factor. The model is built so that the gateway effect is precluded and only a common variable, "drug use propensity", that influences the use of each substance at each stage, is allowed for. They test the hypothesis by using a simulation technique and find that no gateway effect is needed to obtain the generally observed pattern. Morral and colleagues emphasise that their results do not disprove the existence of a gateway effect only demonstrate a possible alternative.

The core question is whether the sequential initiation pattern of drug use is best explained by the mechanisms substantiating the gateway hypothesis or whether the phenomenon is better understood by employing the concepts of accessibility and/or transition proneness. The answer should be of great importance for policy makers and others dealing with the prevention of drug problems. Empirical testing of the gateway theory, i.e. examining whether the observed phenomenon is due to correlation or causality, is, however, inherently difficult. As shown below, analysts have approached the identification problem in different ways and employed various estimations methods.

Empirical testing of the gateway hypothesis

The empirical literature shows contrasting results regarding a possible gateway effect. Some researchers report a strong and significant influence of previous drug consumption on current consumption of the drug, while others present results that do not support the gateway hypothesis. There are at least two possible explanations for the divergence in results. First of all, differences in data could obviously lead to different conclusions. When, for instance, van Ours (2003) concludes that data from the Netherlands in some cases support the gateway hypothesis this need not contradict Pudney's (2003) finding of only a very small gateway effect in a sample of British youths. To some extent the differences may be caused by different rates of response, various sample selection criteria (e.g. age groups), the timing of the survey and a host of other rather mundane but still important factors that lead the researchers to different conclusions. Secondly, different results may be caused by different approaches and methods employed when analyzing the data. For example, different distributional assumptions give ample scope for two researchers with identical data to reach different conclusions. Although important, it is often difficult to test conjectures about the more practical data problems. The problems of how to best approach the data, however, can be explored more theoretically.

Testing the gateway hypothesis illustrates the classical problem of separating heterogeneity and causal effects. Simply documenting that most heavy drug users started with legal drugs and cannabis is not sufficient to establish a causal link. The problem can be illustrated as follows: assume that the probability of starting with e.g. amphetamine is estimated by ordinary regression analysis on the following equation:

$$(1) \quad h_{it} = \alpha_t + \beta X_{it} + \delta d_{it-1} + \varepsilon_{it}$$

in which h_{it} is the risk of starting with amphetamine for a person (i) at a point in time (t); X_{it} is a vector of exogenous variables influencing the probability other than previous use of other drugs (e.g. gender, childhood experiences, peer influence and so on), d_{it-1} is a dummy representing previous use of other drugs and ε_{it} is the error term. In this context h_{it} is the “outcome” of interest and d_{it-1} is the “treatment” to be evaluated. If the dummy for previous drug use turns out to be statistically significant, it's tempting to conclude that the gateway hypothesis is supported by the data. The problem, however, is that standard regression analysis on the equation will produce misleading results if potentially important variables that could explain amphetamine use are omitted. Moreover, some of these omitted variables may influence not only the probability of amphetamine initiation, but also initiation of other drugs.

In this case we will get biased estimates because the dummy variable for previous drug use will capture not only the “true” gateway effect, but also the effect of the omitted variables.

One obvious way to reduce the problem would be to include more variables in the model, on the assumption that selection into the different treatment regimes – in this case past drug use – is ignorable, after conditioning on all of these observable covariates. Yamaguch and Kandel (1984) and Fergusson and Horwood (2000) are two examples of studies that have included a wide range of variables assumed to influence drug use and deviant behaviour. A problem with this approach, however, is that one will never be sure that every relevant variable actually is included. This need not be due to ignorance on behalf of the researcher, but may be caused by lack of data or inherent problems in measuring some potentially important variables. Although the probability is reduced with more variables taken account of, the possibility for a spurious gateway effect still remains and the estimate of the casual effect may be biased by selection on unobservables.

Two possible approaches can be adopted to overcome the problem of selection on unobservables. Firstly, one can employ an instrumental variable (IV) technique that predicts the d_{it-1} on the basis of another variable, or a vector of variables, that are highly correlated to previous drug use but not to the error term in (1). The approach has been adopted by Pacula (1998) who uses past prices of alcohol as instruments for previous consumption of the drug and estimates a gateway effect of alcohol on current marijuana use. She uses data from the National Survey of Youth (NLSY) and reports that higher past alcohol prices are associated with lower likelihood of using marijuana. The same data set, but covering different years, is employed by DeSimone (1998) who uses information on individual characteristics and local prices as instruments. Also Beenstock and Rahav (2002) use variants of the IV approach to sequences of events when they employ prices by birth cohorts as instruments. The main problem with the IV-approach is finding good instruments. Alcohol and cigarette prices have been frequently used. They vary over time and between countries and states, but they cannot reflect contemporaneous individual differences in behaviour within the same area. Prices of illicit drugs are, in addition, hard to obtain. Credible instruments for previous consumption that are not based on prices are rare.

As a second alternative, one can employ models that take account of an unobserved factor that is possibly influencing both the dependent variable and d_{it-1} . Some analysts have taken as their starting point that no study, despite survey method or level of details, will

manage to catch every relevant variation in personal characteristics, experience and environmental factors. They therefore have adopted techniques to account for the unobserved heterogeneity and by that eliminate the spurious effects. The approach relies on the existence of stable and persistent "individual effects" that influence a range of behaviours. Two recent studies have adopted this approach and employed transition models (time-to-event) for estimation. van Ours (2003), combining three surveys from Amsterdam, estimates a bivariate duration model in which he simultaneously analyses the initiation of cannabis and cocaine use. By taking account of correlation in the error terms for the two drugs he is able to control for unobserved heterogeneity using a finite mixture estimator. The work calls on an important identification result by Abbring and van den Berg (2003) in which it is demonstrated that, given appropriate assumptions, casual effects can be identified in bivariate duration models without the need for exclusion restrictions. Van Ours reports that, although the parameter indicating a gateway effect of cannabis on cocaine use is substantially reduced when heterogeneity is controlled, some effect remains, suggesting that there is still a causal link between the two drugs.

Pudney (2003) uses a similar approach to van Ours but he employs British micro data on both drugs and crime. The approach differs in that Pudney uses a discrete time model and a parametric specification of the unobserved heterogeneity, with models estimated by maximum simulated likelihood estimation. In addition, he deals with a broader range of problematic behaviour (solvent abuse, soft drug use, "social" drug use, hard drug use, minor offending and serious crime). After allowing for unobserved individual heterogeneity, however, Pudney reports that the estimates of the dynamic impacts are reduced considerably and remain small even where statistically significant. Consequently, he concludes that the British Government's re-classification of cannabis most probably will not lead to an increase in the number of hard drug users and suggests, on basis of the other variables included, that an effective policy directed at reducing the extent of social deprivation might have a better chance of success.

3 The Oslo study

We employ the unobserved heterogeneity approach by taking account of common unobservables when examining whether data collected among young adults in Oslo shows evidence of a gateway effect from the use of cannabis to a subsequent use of amphetamine and later to cocaine.

3.1 Methods

We proceed in three steps: The first step is to examine whether the individuals in our sample start to use alcohol and illicit substances according to the gateway hypothesis, i.e. examine whether they start with alcohol before cannabis, cannabis before amphetamine and so on. The section employs tools from survival analysis and we estimate separately for each drug the probability of starting to use the drug at different ages given that they haven't used the drug previously (i.e using the Kaplan-Meier method to estimate hazard functions).

Second, we will estimate three separate single equation probit models to determine the statistical relationship between problematic (frequent) use of each substance (cannabis, amphetamine and cocaine) and the following independent variables: gender, social problems (problems with, parents, school, friends and police), attitudes towards free cannabis sale and previous use of other drugs. Previous drug use is included as a dummy. The three models serve as benchmarks against which we can judge the results that allow for unobservable heterogeneity. In line with the gateway hypothesis we have adopted the view that people first start with alcohol, then some proceed to cannabis and later on start with amphetamine and cocaine. This means that we test for previous alcohol use when we estimate the probability of starting with cannabis, only test for previous alcohol and cannabis use when estimating the probability for amphetamine use and test for alcohol, cannabis and amphetamine when cocaine is used as a dependent variable. As use of ecstasy had low prevalence when people in the oldest age groups were in the typical age for experimenting with this drug we have not included ecstasy in the analyses.

In the third step we estimate the three equations together using a multivariate probit specification. This model has been characterized as an "unfairly neglected procedure" (Lesaffre and Molenberghs 1991) in the context of medical statistics and a search of the economics literature indicates that it is no less true of economics. As mentioned previously, we may suspect that the single equations omit relevant variables which we may interpret as "unobserved heterogeneity". The effect of this heterogeneity is captured by the error term in the single equations. The idea behind the multivariate probit model is to model the correlation between the error terms from the single equation models. If there is a systematic relationship between these, one may conclude that an important variable that affects all of the equations has been left out. One may then exploit this systematic relationship between the error terms in

the different equations to allow for the unobserved heterogeneity. Thus, by estimating all three equations at the same time, taking account of the cross-correlation in the error terms, one reduces the problems of unobserved heterogeneity which is a major problem when testing the gateway hypothesis (see Greene (2002) for more on estimation of the multivariate probit model and Contoyannis and Jones (2004) for a recent application that uses the multivariate probit model to estimate a recursive system similar to the one used here). The final element is to compare the results from the single probit models with the multivariate in addition to testing for potential problems. By comparing the results from the two steps we may examine the extent to which correcting for unobserved heterogeneity affects the sign and statistical significance of the estimated gateway effects.

3.2 Data

The data were collected through postal questionnaires sent to a representative sample of 21-31 year olds living in Oslo in 2002. The response rate was roughly 50 per cent with more women than men answering the questions (see Table 1). Only one reminder was sent and a total of 4561 questionnaires were registered. The respondents reported their experience with licit and illicit drugs in addition to socio-economic information on age, gender, education, income and possible childhood problems with parents, friends, school and police, and their attitude toward free sale of cannabis.

(Table 1 about here)

As mentioned in the introduction, we focus on "users" in this study and employ a dummy based on frequency of drug use as the dependent variable for each drug in question. The frequency variables are set equal to one if the respondents report to ever have used the drug more than 25 times. Out of the 40 per cent reporting to have tried cannabis, about one third (13%) have used the drug on more than 25 occasions. The corresponding numbers for amphetamine are 11 per cent having ever tried the drug and 3 per cent are regular users. Ten per cent in the sample report to have ever tried cocaine while 2 per cent have used it on a regular basis. The percentage having ever used alcohol is high (93%) and only 1.5 per cent have ever tried heroin.

Based on a certain set of birth dates for the years 1972-1981 the sample was drawn from the national register. Larger birth cohorts in the first part of the 1970s and a higher

response rate among the older respondents have resulted in a relatively larger fraction from the oldest age groups. The average age was almost 27 years. Less than ten per cent reported problems during childhood with parents (8%), school (6%), friends (4%) and police (1.5%), respectively. In line with the pattern for initiation found in other studies, the average debut age suggest that drug consumers in Oslo start to use alcohol prior to cannabis, then proceed to amphetamine and cocaine. The use of cocaine has, on average, the highest debut age.

"Freesale" is a dummy set equal to one if the person has ticked for being in favour of free sale of cannabis. One may assume that this variable expresses the respondents' underlying attitude towards drugs in general, and as such, one may expect it to be highly positively correlated to their drug use. On the other hand, previous drug use could influence the individual's current opinion about free cannabis sale and we do not know the respondents original attitude to the question, i.e. we do not know which came the first; the attitude or the behaviour. The data reveals that of the 16 per cent that are in favour of free cannabis sale a large fraction of them (78%) report to have consumed the substance themselves. That means, however, that far from every cannabis user favours free sale and not every person supporting free sale has smoked the drug. Despite the possibility that changed attitudes may follow behaviour in this case, we have chosen to include the dummy variable.

The Norwegian Institute for Alcohol and Drug Research has conducted annual surveys among 15-20 year olds in Oslo since 1968. The prevalence variables (cannabisprev, amphetprev and cocaineprev) are based on the general prevalences reported by the youngsters for each of the three drugs in the year the present sample's participants were at the average age for starting with the drug (for example, for a person being 24 in 2002, we used the general prevalence of cannabis in 1997, when they were 19), the general amphetamine prevalence in 1998 (when they were 20) and the cocaine prevalence in 2000 (when they were 22)). Although the prevalences among 15-20 year olds are lower than the corresponding numbers in the current age groups, they may serve as indicators for the cultural and physical accessibility of each drug in question. Only the prevalence variable for cannabis is used in the cannabis equations, only the amphetamine prevalence in the amphetamine equations and so on.

The explanatory variables with the highest interest attached are, however, those representing previous drug use (or "treatment" = d_{it-1} in (1)). Clearly, if there are certain pathways to heavy drug use, one drug must be used before another and we have constructed dummies for the possible gateway drugs. In each equation the dummies are constructed to be

positive only when the individual has used a drug *before* the other drug. Hence, in the equation for amphetamine, the dummy for cannabis is zero for those who have never used cannabis and for those who have used amphetamine before cannabis since in that case cannabis could not be a gateway for amphetamine. Similarly, the dummies in all the equations were constructed to capture only whether the gateway drug had been used before the drug under consideration.

The data's representativeness is hard to assess. It is well known that in general surveys like the one used here, homeless and institutionalised people are under-represented as are people with many sorts of deviant behaviour. This selectivity is especially worrying when the topic of interest is illicit drug use and one may, perhaps, assume that the sample is representative only for the "normal", well-functioning fraction of the population. Reported income and educational achievements suggest that the sample is better off than the average of young people in Oslo. Still, the relatively high prevalence of illicit drug use in the present sample indicates that a large proportion of drug users respond to postal questionnaires.

Recall bias may be another problem, especially here where people are asked to recall the debut age of incidents that occurred, in some cases, more than a decade before. One may argue, however, that using an illicit drug for the first time is so unique that users will tend to remember it. In line with this, one recent study of response reliability in adolescent substance use progression suggests that the reported sequences were reported consistently when checked again three years after the first interview (Golub *et al.* 2000).

3.3 Results

To get a first impression of whether there is a gateway effect, it is useful to explore the order in which people have used various substances (see Table 2). The table confirms the general impression that "soft" drugs are used before "hard" drugs. For instance 10.9% of the total sample claims to have used cannabis before the other illicit drugs (amphetamine, cocaine, ecstasy and heroin), while only 1.5% claimed to have used one of these drugs without using cannabis first. Among the 503 amphetamine users in the sample only 77 per cent report to have first used cannabis and 14 per cent started to use both substances within the same year. The corresponding numbers for the 459 cocaine users were 89 and 7 per cent.

(Table 2 about here)

In addition to giving an impression of the sequences of drug use that are most common, the table also helps to suggest which sequences it is worth testing for in the regression analysis. For instance, very few individuals used cocaine before cannabis and for this reason we have not included a dummy to test whether cocaine could be a stepping stone to cannabis use. Instead we have focused on the major pathways and the table indicates that the most common “stepping stone” is cannabis to some of the other drugs (amphetamine and cocaine).

The "staircase" in drug use initiation is illustrated in Figure 1 where the highest hazard rate for starting with alcohol peaks at an earlier age than the highest hazard rate for cannabis and use of amphetamine and cocaine. The hazard rates give the probabilities for various age groups of starting with a drug given that the person has not started up to that age.

(Figure 1 about here)

We also checked the hazard rate for heroin (n=67), and found that it deviates from that of the other substances by having a less uniform pattern with one peak corresponding to the age of 20 and one at the age of 22, but the small sample size is problematic

Univariate probit models

In Table 3 we present the results of separate estimates of univariate probit models for cannabis, amphetamine and cocaine. The dependent variables in these regressions are not whether the individual has used a substance, but whether the individual has used a substance frequently or not. As argued in the introduction, this is the most policy relevant variable since the justification for making softer drugs illegal is based on the dangers of developing a problematic use of another substance. By including dummies for previous use of drugs further down the staircase provide for preliminary evidence of possible gateway effects. Due to the nonlinearity of the probit function we have also calculated and displayed the partial effects for each of the estimations. These are based on the sample mean values of the regressors and indicate the absolute change in probabilities that occurs when the variable of interest changes by one unit (continuous variables) or when a dummy variable changes from zero to one in value. Unless otherwise stated, the parameters are statistically significant at a 5 per cent level.

(Table 3 about here)

As shown in Table 3, there seems to be an influence of previous alcohol use on later frequent cannabis use, i.e. there seem to be a gateway effect of alcohol to cannabis. Given average values of the other explanatory variables the probability of frequent cannabis use increases by 0.07 if the individual has used alcohol previously. Males are more prone than females to be frequent cannabis users as are people that report they have had problems with parents, school and police during their childhood. Problems with police have a stronger influence on regular cannabis use than have the other two problem factors. In contrast to this, people reporting to have had childhood problems with friends (significant at a 10 per cent level), are less likely to be regular cannabis users, something which may confirm that cannabis smoking is a social activity. Further, the results indicate that those who favour free sale of cannabis are more likely to use cannabis frequently. In fact, the partial effect of "freesale" suggests an increase in probability of 0.30 for a person in favour of free cannabis sale compared to an identical individual who is against such sale. The cannabisprev variable, assumed to be an indicator for physical and cultural accessibility of the drug, suggests a positive relationship between the general prevalence of cannabis and the probability of frequent cannabis use.

Further, looking at the results for the regular use of amphetamine in Table 3, we find a substantial influence of previous cannabis use on subsequent use of amphetamine while the effect for alcohol is smaller and statistically insignificant. Also the parameters for gender and childhood problems with friends are insignificant in a statistical sense. The other types of childhood problems, however, are positively correlated to regular amphetamine use with police problems being more important than problems with school and parents. The results also show that the attitude towards free cannabis sale has a smaller impact on frequent amphetamine use than it had on the use of cannabis. And similar to the results for the cannabis equation, the prevalence variable, here measuring general amphetamine use at the time when the respondents were aged 20, is positively correlated to frequent use of the drug in question. As the overall probability of frequent amphetamine use is relatively small (3 per cent), so are the calculated marginal effects in this case. As can be seen from Table 3, childhood problems with police and previous cannabis use have the largest effects and both increase the probability of frequent amphetamine use by 0.06.

The results for the cocaine equation reveal that both previous cannabis and amphetamine use are positively associated with for regular use of cocaine. The alcohol dummy, however, is negative and insignificant. There is no significant difference between the genders and, of the variables indicating childhood problems, only problems with the police obtain a significant estimate. The latter variable has a smaller influence on regular cocaine use than the variable had on regular cannabis and amphetamine use. The data does not indicate any statistically significant effect of the prevalence variable nor of the respondents' attitudes towards free sale of cannabis.

Multivariate probit models

The interesting question now, however, is whether the substantial gateway effects found in the separate estimation of the three equations remain after we have taken account of unobserved heterogeneity. The results from the multivariate probit model are reported in Table 4.

(Table 4 about here)

The most striking result is the reduced values of the gateway effects in the equations for amphetamine and cocaine: For amphetamine the coefficient on the cannabis dummy has changed from 1.45 to 0.70 and for cocaine from 1.10 to a statistically insignificant value of 0.63. Further, the coefficient on the amphetamine dummy in the cocaine equation decreases from 1.56 to 0.51. The alcohol dummy remains insignificant for both substances. The sizeable correlation coefficients for the three equations are presented at the bottom of Table 4 and indicate the importance of estimating the equations as a system. All of the correlations are positive, consistent with the idea of a common unobservable propensity to substance abuse. In Table 4 the value of the freesale parameter for amphetamine is higher compared to the corresponding value in Table 3 whereas the estimates for the other variables in the amphetamine equation remain fairly unchanged. When comparing the univariate and the multivariate probit results for cocaine, we see that more coefficients are significant at a 10 per cent level in the multivariate probit model (parents and school). In addition, the estimates for childhood problems with parents, school and police and the freesale variable have increased values in Table 4. In contrast, there is hardly any difference between the parameter values in Table 3 and Table 4 for the cannabis equation, which means that taking account of

unobserved heterogeneity has not had any influence on the estimated gateway effect or on the other explanatory variables for this drug.

Table 5 presents the partial effects based on the coefficients from the multivariate probit model. They are computed at the sample means of the regressors for each of the three substances separately. For the dummy variables the partial effects show the difference in predicted probability of becoming a regular drug user when the dummy is 1 or 0. As the overall probabilities of frequent amphetamine and cocaine use are relatively small, we have also included the percentage changes in these predicted probabilities. The intention is to underline the quantitative importance of the various dummy variables. Statistically significant values are in bold, and we can see that the dummies for childhood problems with school and police, as well as the respondents' attitude towards free cannabis sale, are associated with a substantial change in the probabilities for frequent drug use across the three substances. For instance, given identical values of the other variables the probability of becoming a frequent amphetamine or cocaine user increases by 87 per cent in both cases if the individual has had problems with the police during childhood. The absolute partial effects are 0.08 and 0.02, respectively. Also the percentage increase in the probability of previous cannabis use on later amphetamine use and previous amphetamine use on later cocaine use are substantial. Hence, these variables are potentially of great importance even though the overall probability of becoming a frequent user of these drugs is relatively small.

In order to test sensitivity of the multivariate probit results reported in Tables 4 and 5, we re-ran the model with the cut-off point for frequency of use set to 11 occasions or more (previously it was set at 25 occasions). This means that more people were included as "frequent users". For cannabis, 225 people (12.4 percent of those reporting cannabis use) changed status, whereas the corresponding numbers for amphetamine and cocaine were 72 (14.3 per cent of all amphetamine users) and 69 persons (15.0 per cent of all cocaine users). The multivariate probit results with these new dependent variables were fairly similar to those presented above. The same set of coefficients were statistically significant and the signs were retained. The difference between the parameter values were not systematically positive or negative and not substantial for the statistically significant estimates. One may therefore infer from this that because many people report that they stop taking drugs after they have tested it once or twice, the cut off point of 25 occasions or more of using a substance seems not to influence the results substantially.

4 Discussion and Conclusions

Drug policy and legislation in many countries have been influenced by the gateway theory. The influence is especially apparent in the case of cannabis, often called "the gateway drug". The common finding that people having consumed cannabis have a higher probability of using hard drugs (see for instance Kandel et al. 1992 for a list of such studies) has been taken as evidence to support the theory. Comparisons of cannabis users versus non-users have clearly shown that users to a larger degree try other illicit drugs and more so if they are frequent cannabis users (Ferguson and Horwood 2000). It is important to remember, however, that both temporal precedence and correlation are necessary although not sufficient to establish causality. The correlation may be spurious reflecting some third factor influencing the use of several types of drugs.

In the present paper we have analyzed data collected in 2002 from a representative sample of 21-31 year olds in Oslo. In contrast to comparable studies just focusing on whether the respondents have tried a drug or not, our object has been examining a possible gateway effect on *regular* drug use, as most drug problems are related to frequent users of drugs. Further, to reduce the possibility of omitted variable problems we have included several potentially important variables that could influence frequent drug use and we have employed estimation methods that also take account of unobserved heterogeneity when we estimated the gateway effect.

Estimates of the gateway effects and the additional explanatory variables are presented in Tables 4 and 5. Despite the common finding that males are over-represented among problem drug users, there were no gender differences in the probability of becoming a frequent amphetamine or cocaine user in the current sample. Gender seems to be important only for frequent cannabis use with males having a higher probability of becoming regular users than females. According to the hypothesis of "transition proneness" drug use, truancy and various types of crimes etc may all express vulnerability to deviant behaviour resulting from environmental conditions or personal characteristics. Self-reported problems during childhood were included to account for this. Problem with parents, school and police all seem to be of importance for regular use of the three substances. Problems with the police increased the probability of frequent drug use more than problems with parents and school. Whether

people reported to have had problems with friends, on the other hand, did not influence their probability of becoming frequent drug users.

"Freesale", measuring whether the respondents favour free sale of cannabis, was another type of variable included. As mentioned, we expected it to be highly correlated to own drug use as it may be interpreted as an indicator of attitudes towards drugs, although not every person claiming to be in favour of cannabis legalisation reported to have used illicit drugs themselves. The estimates show that the variable has a significant influence on frequent use of the three drugs with more influence on frequent cannabis use than on the other two. Although we cannot exclude the possibility that drug use itself subsequently changed individuals' attitudes, the result may suggest that differences in personality and attitudes play an important role in regular drug use.

The last set of variables we included in the models was measures of drug use prevalences taken from annual survey data of 15-20 year olds in Oslo. We assumed that the percentage of youngsters reporting to have tried e.g. cannabis in a given year could be taken as an indicator of the general physical and cultural accessibility of the drug. Consequently, the accessibility was assumed to be generally higher for those aged 19 years in 2000 when the cannabis prevalence was 28.6 per cent than for those who were 19 in 1991 when the prevalence was 16.6 per cent. Only the cocaine prevalence employed in the cocaine equation did not seem to have a significant influence on regular use of the drug. This could be due to cocaine use being less common among 15-20 year olds and thereby the prevalence for this particular drug being a poorer indicator for cocaine use in general.

Including the above mentioned variables has proved important, but as the survey does not provide us with all the variables that could possibly influence frequent drug use, we have employed methods that take account of unobservable individual-specific effects to reduce the possibility of estimating a spurious gateway effect. The multivariate probit model is presented in section 3. The results shown in Tables 4 and 5 suggest that there *are* some gateway effects of previous drug use on subsequent frequent use of cannabis, amphetamine and cocaine, respectively. The effect of previous drug use was greater in estimates where we did not take account of unobserved heterogeneity (Table 3), but, although substantially reduced, they remain substantial when unobserved factors are accounted for. The tendency of reduced influence of the gateway variables after taking account of unobserved heterogeneity is in line with the findings of van Ours (2003) and Pudney (2003). They employ alternative methods

for taking unobservables into account. According to the current findings, alcohol is a gateway drug for cannabis, cannabis is a gateway drug for amphetamine and amphetamine is a gateway drug for cocaine.

Given that there is an effect of previous drug use on subsequent regular use of a drug further down the pathway to hard drug use, what are the policy implications of the finding? Should we ban alcohol to reduce later cannabis problems and will the recent policy change towards cannabis in the UK leads to an increased number of amphetamine users in the next few years? As discussed in section 2, there are at least five possible mechanisms that may explain an observed gateway effect and what drug policy to recommend will, among other things, depend on which of the mechanisms that actually operates. If the use of one drug creates a psychological or physiological need for further and stronger experiences of the same type, if the use on one drug reduces the costs of starting with another and more dangerous drug or if the use of one drug increases the utility of consuming another, then a strict drug policy may be the preferred option. An extensive cost-benefit analysis including various aspects of such a restrictive policy, however, is needed to determine the question.

On the other hand, if the act of obtaining a soft drug brings the user into contact with hard-drug users or suppliers whom he/she would not otherwise have met, then the Dutch option may seem more attractive. Separating the markets for soft and hard drugs by legalising consumption and sale of cannabis may then prove successful in reducing the rate of transition to hard drug use. Further, if people, after experiencing no obvious ill effects of soft drug use, have reduced confidence also in the strong negative publicity directed against hard drug use, the solution may be to make more distinct the differences between the various drugs, perhaps including more distinct differences in the legal sanctions directed against drug use and trade.

Acknowledgements

The authors would like to thank participants at the 24th Arne Ryde Symposium: The economics of substance use, Lund 13-14 August 2004 for useful comments to an earlier draft of the paper.

References

- Abbring, J.H., van den Berg, G.J., 2003. The nonparametric identification of treatment effects in duration models. *Econometrica* 71(5), 1491--1517.
- Becker, G.S., Murphy, K.M., 1988. A Theory of Rational Addiction. *Journal of Political Economy* 96, 675--700.
- Beenstock, M., Rahav, G., 2002. Testing Gateway Theory: do cigarette prices affect illicit drug use? *Journal of Health Economics* 21, 679--698.
- Contoyannis, P., Jones, A.M., 2004. Socio-economic status, health and lifestyle. *Journal of Health Economics* 23, 965--995.
- DeSimone, J., 1998. Is marijuana a gateway drug? *Eastern Economic Journal* 24, 149--164.
- Ferguson, D.M., Horwood, L.J., 2000. Does cannabis use encourage other forms of illicit drug use? *Addiction* 95, 505--520.
- Golub A., Labouvie E., Johnson B.D., 2000. Response reliability and the study of adolescent substance use progression. *Journal of Drug Issues* 30(1), 103--118.
- Goode, E., 1972. Cigarette smoking and drug use on a college campus. *The International Journal of the Addictions* 7(1), 133--140.
- Greene, W.H., 2002. *Econometric Analysis* (fifth edition). Prentice Hall, New Jersey.
- Jessor, R., Chase, J.A., Donovan, J.E., 1980. Psychosocial correlations of Marijuana use and problem drinking in a national sample of adolescents. *American Journal of Public Health* 70 (6), 604--613.
- Johnston, L. D., 1991. Toward a Theory of Drug Epidemics, in: Donohew, L., Sypher, H.E., Bukoski, W.J. (Eds.), *Persuasive Communication and Drug Abuse Prevention*, Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 93--131.
- Kandel, D.B., 1975. Stages in adolescent involvement in drug use. *Science* 190, 912--914.
- Kandel, D.B., Yamaguchi, K., Chen, K., 1992. Stages of progression in drug involvement for adolescence to adulthood: further evidence for the gateway theory. *Journal of Studies on Alcohol* 53 (5), 447--457.
- Kenkel, D., Mathios A.D., Pacula R.L., 2001. Economics of youth drug use, addiction and gateway effects. *Addiction* 96, 151--164.
- Lesaffre E., Molenberghs, G., 1991. Multivariate probit analysis: A neglected procedure in medical statistics. *Statistics in medicine* 10(9), 1391--1403.
- MacCoun R.J., Reuter P., 2001. *Drug war heresies: learning from other vices, times, and places*. Cambridge, U.K.: Cambridge University Press; New York.
- Morrall, A.R., McCaffrey, D.F., Paddock, S.M., 2002. Reassessing the marijuana gateway effect. *Addiction* 97, 1493--1504.
- Pacula, R.L., 1997. Economic modelling of the gateway effect. *Health Economics* 6(5), 521--524.

Pacula, R.L., 1998. Does increasing beer tax reduce marijuana consumption? *Journal of Health Economics* 17, 557--585.

Pudney, S., 2003. The road to ruin? Sequences of initiation to drug use and crime in Britain. *Economic Journal* 113, 182--198.

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

Yamaguchi, K, Kandel, D.B., 1984. Patterns of drug use from adolescence to young adulthood: III. Predictors of progression. *American Journal of Public Health* 74, 673--681.

Table 1. Description and definition of variables (n=4561)

	Means and standard deviations	Definitions
Cannabisfreq	0.13 (0.33)	Dummy; 1 if used more than 25 times
Amphetfreq	0.03 (0.18)	Dummy; 1 if used more than 25 times
Cocainefreq	0.02 (0.14)	Dummy; 1 if used more than 25 times
Gender	0.41 (0.49)	Dummy; 1 if male
Age	26.6 (2.63)	Age in years
Agedummy	0.33 (0.47)	Dummy; 1 if younger than 26
Parents	0.08 (0.27)	Dummy; 1 if problems in childhood
School	0.06 (0.24)	Dummy; 1 if problems in childhood
Friends	0.04 (0.19)	Dummy; 1 if problems in childhood
Police	0.015 (0.12)	Dummy; 1 if problems in childhood
Freesale	0.16 (0.37)	Dummy; 1 if support free cannabis sale
Cannabisprev	20.8 (3.64)	Annual prevalence rate at respondents' age of 19
Amphetprev	4.88 (1.81)	Annual prevalence rate at respondents' age of 20
Cocaineprev	3.21 (1.36)	Annual prevalence rate at respondents' age of 22
Cannabismin	19.1 (3.41)	Debut age for using cannabis (n=1807)
Amphetmin	20.1 (3.37)	Debut age for using amphetamine (n=503)
Cocainemin	21.7 (3.41)	Debut age for using cocaine (n=463)

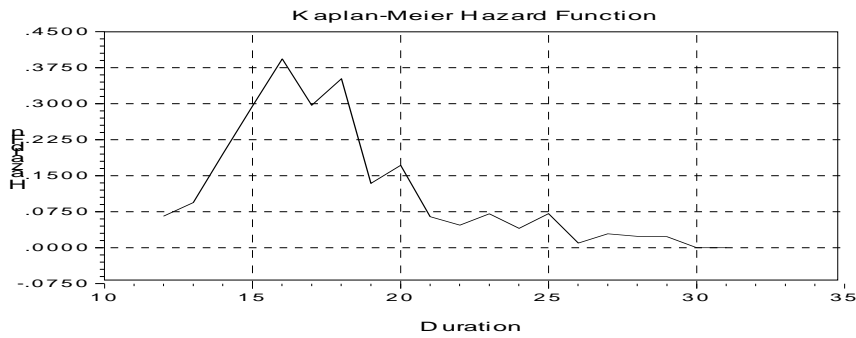
Table 2 Paths of substance use in the sample (n=4561)

Description	Percentage
No drugs or alcohol	4.5
Only alcohol	51.5
Only cannabis	2.4
Alcohol → Cannabis (only)	24.9
Cannabis → Alcohol (only)	0.5
Cannabis → Amphetamine/Cocaine/Ecstasy/Heroin	10.9
Amphetamine/Cocaine/Ecstasy/Heroin without previous cannabis use	1.5

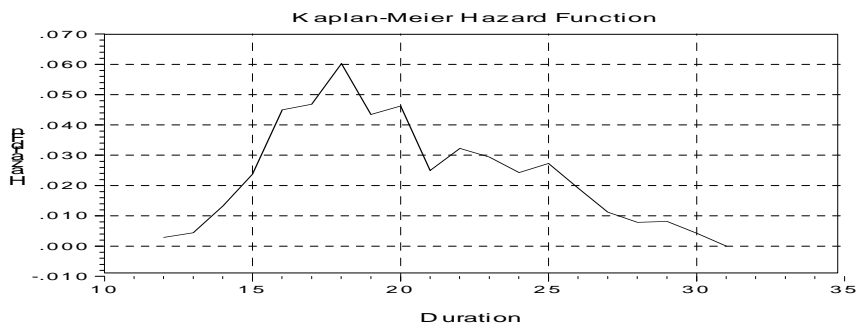
* The results are based on whether the individual had ever used one of the drugs and at what age.

Figure 1. Hazard rates for onset of alcohol, cannabis, amphetamine and cocaine use.

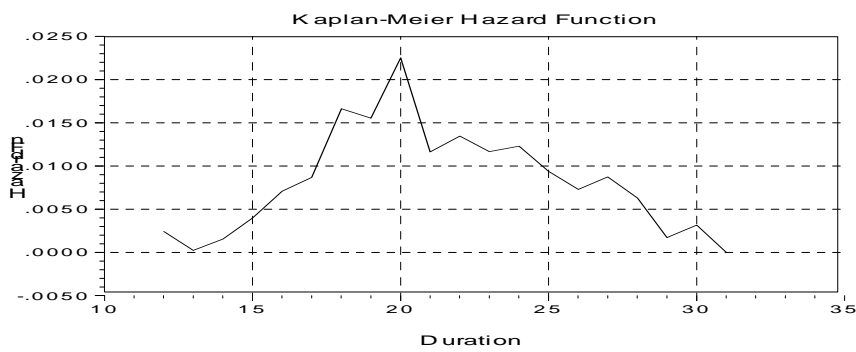
Alcohol



Cannabis



Amphetamine



Cocaine

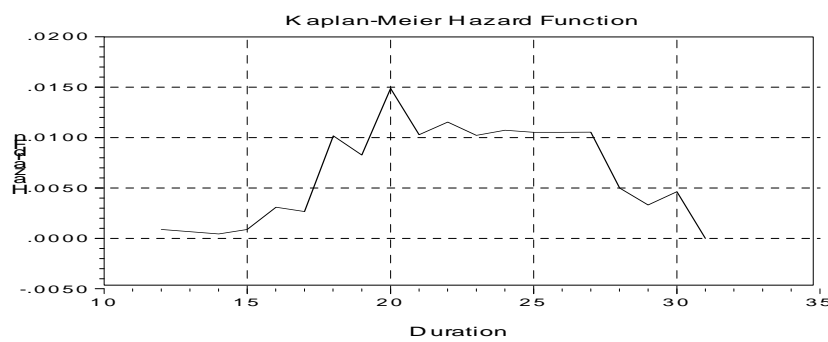


Table 3. Coefficients and partial effects from univariate probit models (standard errors in parentheses). N=4561. Dependent variables: Dummy for frequent use (>25 times) of cannabis, amphetamine and cocaine, respectively. Intercepts not reported.

	Cannabis		Amphet.		Cocaine	
	Coefficient	Partial effect	Coefficient	Partial effect	Coefficient	Partial effect
Gender	0.22 (0.05)	0.04 (0.009)	0.03 (0.09)	0.001 (0.002)	0.06 (0.13)	0.0002 (0.0005)
Parents	0.55 (0.09)	0.12 (0.02)	0.23 (0.12)	0.006 (0.004)	0.20 (0.17)	0.0009 (0.001)
School	0.39 (0.10)	0.08 (0.03)	0.52 (0.13)	0.02 (0.009)	0.21 (0.18)	0.0009 (0.001)
Friends	-0.26 (0.15)	-0.04 (0.02)	-0.01 (0.19)	0.0002 (0.004)	0.06 (0.26)	0.0002 (0.001)
Police	0.86 (0.18)	0.23 (0.06)	0.91 (0.19)	0.06 (0.025)	0.48 (0.24)	0.003 (0.004)
Freesale	1.16 (0.06)	0.30 (0.02)	0.35 (0.09)	0.01 (0.004)	0.17 (0.13)	0.0007 (0.0008)
Alcoholcum	0.54 (0.14)	0.07 (0.01)	0.04 (0.34)	0.001 (0.007)	-0.40 (0.55)	-0.002 (0.005)
Cannabisprev	0.04 (0.01)	0.01 (0.001)				
Cannabiscum			1.45 (0.17)	0.06 (0.007)	1.10 (0.39)	0.007 (0.003)
Amphetprev			0.05 (0.02)	0.001 (0.001)		
Amphetcum					1.56 (0.15)	0.045 (0.02)
Cocaineprev					0.03 (0.05)	0.0001 (0.0002)
Loglikelihood	-1393		-506		-260	

Table 4. Coefficients from multivariate probit model. n=4561. Dependent variables: Dummy for frequent use (>25 times) of cannabis, amphetamine and cocaine, respectively. Intercepts not reported.

	Cannabis	Amphet.	Cocaine
Gender	0.22 (0.05)	0.02 (0.09)	0.05 (0.12)
Parents	0.54 (0.09)	0.29 (0.12)	0.29 (0.15)
School	0.40 (0.11)	0.56 (0.13)	0.42 (0.20)
Friends	-0.25 (0.15)	-0.03 (0.20)	0.004 (0.28)
Police	0.86 (0.20)	0.92 (0.21)	0.73 (0.25)
Freesale	1.16 (0.06)	0.57 (0.09)	0.55 (0.13)
Alcoholcum	0.51 (0.13)	0.11 (0.46)	0.01 (0.79)
Cannabisprev	0.04 (0.01)		
Cannabisdum		0.70 (0.16)	0.63 (0.48)
Amphetprev		0.04 (0.02)	
Amphetdum			0.51 (0.16)
Cocaineprev			0.04 (0.04)

Correlation between cannabis and amphetamine	0.70 (0.04)
Correlation between cannabis and cocaine	0.64 (0.07)
Correlation between amphetamine and cocaine	0.78 (0.04)

Table 5. Partial effects from multivariate probit model. n=4561. Dependent variables: Dummy for frequent use (>25 times) of cannabis, amphetamine and cocaine, respectively. Intercepts not reported. Statistically significant coefficients are in boldface (5% level).

	Cannabis		Amphetamine		Cocaine	
	Partial effect	% change	Partial effect	% change	Partial effect	% change
Gender	0.028	35%	0.0006	5%	0.0004	13%
Parents	0.092	62%	0.0117	52%	0.0034	58%
School	0.062	51%	0.0303	74%	0.0060	71%
Friends	-0.026	-69%	-0.0009	-9%	0.00003	1%
Police	0.181	75%	0.0755	87%	0.0169	87%
Freesale	0.246	85%	0.0277	75%	0.0082	80%
Alcoholcum	0.063	64%	0.0031	26%	0.0001	4%
Cannabisprev	0.004					
Cannabiscum			0.0269	83%	0.0069	85%
Amphetprev			0.0013			
Amphetcum					0.0076	77%
Cocaineprev					0.03	