Hard Drug Uptake among Cannabis Users: A Bayesian Analysis

Anne Line Bretteville-Jensen
Norwegian Institute for Alcohol and Drug Research (SIRUS)
email: alb@sirus.no

Liana Jacobi
Department of Economics
University of Melbourne
email: ljacobi@unimelb.edu.au

January 2007

Abstract

The goal of this paper is to analyse the relative importance of the gateway, proneness and accessibility effects in explaining the increased probability of hard drug use among cannabis users. The previous literature has focused on the gateway effect, finding mixed evidence that the use of cannabis increases the probability of using a hard drug later on. While the gateway effect suggests that the initiation pattern commonly found in empirical studies (cannabis use precedes hard drug use) is due to a causal linkage between cannabis and subsequent hard drug use, proneness and accessibility factors are alternative (additional) explanations. This paper uses unique survey data from a representative sample of young adults in Norway, which provides individual indicators for both accessibility and proneness. An interesting feature of this and other drug use data is the small number of hard drug users without previous cannabis use. In our data we almost exclusively observe hard drug use in subjects with previous cannabis use. To address this problem we propose a Bayesian degenerate sample selection model for cannabis use and subsequent hard drug uptake, where cannabis and hard drug intake (only modelled for cannabis users) are modelled jointly to allow for unobserved factors affecting the initiation of cannabis and hard drugs. A Bayesian predictive approach is used to compute the probability of hard drug uptake among cannabis users and to asses the relative importance of the gateway, accessibility and proneness effects in explaining the hard drug uptake. Suitable Markov Chain Monte Carlo methods for the model fitting and the predictive analysis are provided in the paper. Our empirical analysis provides evidence for all three effects. While the magnitude of the effects varies across model specifications, the relative importance remains mostly stable across the specifications.

DRAFT (Do not cite without permission!)

Key words: Drug use, Gateway, Proneness, Accessibility, Markov Chain Monte Carlo; Predictive Analysis

1 Introduction

Cannabis is the most widely-used illicit drug in Europe, Australia and the US and an increasing number of people are seeking treatment for cannabis use (EMCDDA 2006). In the EU countries and Norway prevalence figures are higher now than ten years ago. An interesting and important question is whether the widespread cannabis use will also lead to an increased number of hard drug users. The question emerges because one of the most robust findings in the epidemiology of illicit drug use is the initiation pattern, where most illicit drug users seem to have started with legal drugs like cigarettes and alcohol before proceeding to cannabis and subsequently, if further involvement, to hard drugs (Kandel 2002). Very few hard drug users go straight to heroin without previously having used several other drugs, and cannabis is often the first of the illegal drugs being used. Kandel (1975), employing a US sample, was the first to point out this "staircase" but a similar pattern has been retrieved in many subsequent studies of drug initiation (Kandel 2002).

The observed "staircase" pattern has led to the so called "gateway" hypothesis which states that the use of a soft drug increases the risk of using a hard drug later on, i.e. it claims that there is a causal relationship between the various drugs. An extended version the hypothesis also states that the risk increases if the user starts at a very young age and if the soft drug is used frequently. In many western countries the gateway hypothesis has been influential in formulating the drug policy and opponents of decriminalizing soft drugs (cannabis) often refer to the assumed causal relationship in their argumentation. The possible mechanism(s) that could underpin the causal relationship are still being discussed (Fergusson et al. 2006; Melberg et al. 2007).

There are, however, other possible explanations for the observed higher probability of hard drug use among cannabis users that do no claim a causal relationship between the two. First, hard drugs may be more accessible (economic, cultural and/or physical) to cannabis users than to others. Accessibility to hard drugs may vary to a large extent between youngsters. Some people have better knowledge of how and where to obtain drugs, what prices to pay etc. and drug use is more accepted in some youth cultures than in others. This may imply that cannabis users on average have better access to hard drugs than non-users and explain the higher probability of hard drug use among the first group. Second, a common proneness against deviant behavior may lead some to consume illicit drugs, and cannabis is just used

prior to others because it is cheaper and more readily available. Excessive drug use is then viewed as one out of many possible responses to unfavorable genetic endowments or traumatic childhood experiences. If people with a certain vulnerability have an increased risk for drug use, there may be no causal link between soft and hard drugs that explain the familiar initiation pattern (Morral et al. 2002). In general, the gateway, accessibility and the proneness hypotheses offer different (and co-existing) explanations for the observed correlation between soft and hard drug use, and may have very different policy implications. Empirical studies are needed to improve our understandings of this phenomenon.

The existing literature on the sequential pattern of drug use initiation has mainly focused on the gateway hypothesis, and to a lesser extent on proneness and has had little to say about the influence of accessibility. One reason is that survey data has been frequently employed in the testing and such data do not usually include variables that can account for all three hypotheses. For instance, survey data may not have indicators for drug availability and, generally, will not include information about genetic factors. Recent studies based on retrospective survey data have instead focused on ways of accounting for the possible endogenous gateway variable. Twin and sibling studies and prospective, longitudinal studies of youngsters have, on the other hand, often emphasized proneness in addition the gateway hypothesis but they have not paid much attention to possible effects of varying accessibility. These studies have yielded mixed evidence for the gateway effect. For example, Pacula (1997), DeSimone (1998), Fergusson et al (2006), Melberg et al (2007), Van Ours (2003) and Bretteville-Jensen et al found support for the gateway hypothesis, while Pudney (2003), Beenstock and Rahav (2002) found weak or no support for the gateway hypothesis.

This paper uses a new and unique data set from Norwey to explore the relative importance of the gateway, proneness and accessibility effects in explaining the increased probability of hard drug use among cannabis users. An challenging feature of this data (and other drug use data) and other is that we almost exclusively observe hard drug use in subjects with previous cannabis use. We therefore focus our analysis on the probability of hard drug uptake among cannabis drug users. For any type of data source, examining the gateway effect is inherently challenging as there may be unobserved factors that can influence both the initiation of cannabis and hard drugs. Estimation methods that fail to control for these effects may produce spurious results for the gateway effect. To analyze the effect from cannabis use and

the relative importance of the gateway, proneness and accessibility effects, we propose a "degenerate" sample selection model for cannabis and hard drug uptake that explicitly considers accessibility and proneness. Hard drug intake is only modeled for cannabis users and we account for possible remaining unobserved confounding factors by modeling cannabis and hard drug intake jointly for cannabis drug users. The model is formulated in the Baysian paradigm and estimated via Markov Chain Monte Carlo (MCMC) methods. We then employ the Bayesian predictive approach for the analysis of the probability of later hard drug uptake among cannabis users and to evaluate the relative importance of the gateway, accessibility and proneness effects and explore policy relevant scenarios. Suitable Markov Chain Monte Carlo methods for the estimation of the model and the predictive analysis are derived in the paper. Our results suggest that all three effects contribute to the observed hard drug use pattern.

2 Data

In 1998, 2002 and 2006 the Norwegian institute for alcohol and drug research (SIRUS) sent postal questionnaires to a representative sample of young adults to collect data about the respondents legal and illegal drug use. In this paper we employ a merged file of the later two data sets which contain information on the frequency of current and life time drug use, as well as on the starting ages for the various drugs. In addition, the data contains information on personal characteristics and number of variables that relate to the proneness of a subject's of drug use and his or her accessibility to drugs. The response rate in the two surveys was about 50 per cent which is roughly in line with other studies of this type. Only one reminder was sent to everyone in the sample. The survey was aimed at the population of the 21 to 30 year old. This age group is particularly suited for the analysis of illicit drug use initiation, as the data shows that most recreational drug users will have started by the age of 21 and that frequency and extent of drug use is higher among these young adults than later in life. Also, recall bias may be less of a problem here compared to studies including people in their 40-60s since their drug initiation probably would have occurred several years before the time of survey.

The total number of respondents for which we have information on all the variables of interest is 10,984. The sample procedure used implies an oversampling of people from the capital city, Oslo. This increases the overall prevalence of drug use as young people in Oslo use more drugs than do people from comparable age groups outside the capital but should not otherwise affect the sample. General population

surveys are, however, always associated with methodological problems. One concern is the relatively low response rate and there are reasons to assume that more habitual users of drugs, among them persons who regularly inject drugs, tend to be underrepresented in the net sample. First, a relatively large proportion of this group is rarely found at home. As a consequence, they are hard to reach with the questionnaire. Second, habitual drug users may be less inclined to fill in questionnaires than the average young person. Thus, it is possible that our findings are more relevant for the group of recreational users of illicit and licit drugs than for the group of problematic drug users.

Among the respondents there is also the problem of false negatives (people claiming not to have used illegal substances or reporting a lower consumption level or frequency than the real one) and false positives (people exaggerating their actual drug use). For phenomena with low frequency, like heroin use, the false positive is considered a bigger problem than the false negative (Skog 1992). We have no means to test for false negatives. However, the reported prevalence of the non-existing drug "relevin" gives an indication of false positives in the present sample. "Relevin" was listed as just another drug the respondents were asked about and with only 16 people (0.1 per cent) reporting to have ever used this particular drug, false positives does not seem to be a pervasive problem in this sample.

3 Drug Use Pattern, Proneness and Accessibility of Illicit Drugs in the Norwegian data

3.1 Initiation and Illicit Drug Use Pattern

Before introducing an appropriate estimation framework to investigate the hard drug use in terms of the gateway, proneness and accessibility effects, we examine the initiation pattern in the data. Do young people in Norway start with cannabis before proceeding to harder drugs or has for instance the introduction of ecstasy, that occurred in Norway in the latter part of 1990s, changed the pattern? Figure A1 shows the initiation pattern of drugs by year of survey.

(Figure A1 about here)

The figure clearly shows the commonly reported initiation pattern for drug use. The mean age of starting with alcohol is 15.4 which is much lower than the mean starting ages for illegal drugs (cannabis: 18.6; heroin: 19.5; amphetamine: 19.7; ecstasy: 20.3 and cocaine: 21.1 years). It is somewhat surprising

that the mean age for heroin initiation is lower than comparable numbers for ecstasy and cocaine and but it is not statistically different from the starting age for amphetamine. The low prevalence of heroin (1.4 per cent) might suggest that drug users view this drug differently from the other narcotic drugs and that there is a particular selection to the heroin using group. The figure also shows a decrease in starting ages for cannabis, amphetamine, cocaine and ecstasy from 2002 to 2006. For alcohol and heroin there was no statistically significant decline in starting ages across the periods.

We further examined the "staircase" pattern at the individual level and found that despite a few exceptions and some unsettled cases, the "staircase" seems to apply at this level too. Only 79 people in the sample (0.7% of the sample) had used hard drugs prior to their first cannabis use and 61 people (0.6% of the sample) had only used hard drugs. The first group represents a kind of "reversed gateway", and was therefore excluded from the sample and the further analysis. The second group (hard drug use among non-users of cannabis) could form a comparison group in the statistical analysis when estimating the effect of soft drug use on the probability of hard drug uptake. Unfortunately there are too few observations in this group for this purpose. We therefor exclude these two groups, which reduces the sample size from 10,984 to 10,844. A third group that poses some challenges, are the 2.1 per cent (227 people) that reported the same starting age for cannabis and the hard drugs. Since cannabis use tends to precede hard drug use for most illicit drug users, we have kept these observations in the sample, although though we cannot strictly tell which drug came first for this latter group. To examine the robustness of the results, we will implement the empirical analysis in Section 5 also under the restricted sample that excludes this group.

Table A1 presents the descriptive statistics for the main variables of interest. We have lumped the use of amphetamine, ecstasy, heroin and cocaine together to create the variable "hard drugs", since we do not focus on the possible influence of amphetamine on subsequent cocaine use etc. in this paper. To highlight the differences between groups that vary with respect to illicit drug use, we have divided the sample into a) non-users of any illicit drugs; b) illicit drug users and split the latter group into c) cannabis users only and d) users of both cannabis and hard drugs. For comparison, also the corresponding numbers for the full sample ($n = 10,344$) are displayed.

(Table A1 about here)

More than one third of the full sample reports to have used cannabis at least once and about one in eight has also used a hard drug. The prevalence of cannabis use increased from 35.0 per cent in 2002 to 38.2 per cent in 2006 and the hard drug prevalence from 11.6 to 12.5, yet only the cannabis increase is statistically significant. We see that 16 per cent of the illicit drug users report to have initiated their cannabis use before turning 16 and 1 in 3 report frequent cannabis use. Splitting these numbers for the cannabis-only and hard drug users, Table A1 shows that there are large variations across the groups: hard drug users report early cannabis initiation more frequently (30 vs 9 per cent) and more hard drug users also claim to have used cannabis at least 25 times (63 vs 16 per cent). Thus, these mean values seem to confirm the extended gateway hypothesis about increased risk of hard drug use among young starters and frequent users of cannabis. Table A1 further shows that more females have answered the questionnaire (6 out of 10 are females). Still, more than half of the hard drug users are men (53%) and the lowest male rate is found among non-users of illicit drugs. As expected, there are more soft and hard drug users among young adults in Oslo than outside the capital. The smaller gross (and net) sample in 2006 is due to changes in the sampling procedure.

3.2 Proneness and Accessibility for Illicit Drugs

The questionnaire has several variables that can be taken as indicators of proneness and accessibility. Childhood problems, for example, can be used as indicators for proneness towards drug use. In the survey, the respondents were asked if they had any problems with parents, school, friends and police during their childhood, respectively. Given the mean initiation age for the illicit drugs being above 18 years, it seems reasonable to assume that for most subjects the reported problems refer to a period prior to their uptake of illicit drugs. In line with the proneness hypothesis, Table A1 shows that the dummy variables for childhood problems all have higher prevalence figures among hard drug users than among cannabis-only and non-users of any illicit drug. The difference is larger for problems with parents and school than for problems with friends, yet it is problems with police variable which displays the largest difference across the groups. We see that 30 per cent of hard drug users report childhood problems whereas the corresponding number for cannabis-only and non-users is 15 and 10 per cent, respectively. The large majority of young adults (87%) did not report any of the listed childhood problems.

Table A1 indicates that the group of illicit drug users has a much higher prevalence of respondents

that started with alcohol at a very young age and who dropped out of school after only 10 years. Starting with alcohol before turning 13 could be an indication of a deviant personality or reflect adverse personal experiences. Leaving school at an early age (15/16 years) may be explained by the same reasons or it could suggest a relatively high time preference rate, which according to "the theory of rational addiction" increases the probability of a problematic drug use (Becker and Murphy 1988). The possibility of common, unobserved factors (like the time preferences, childhood traumas etc) that influence both the problem variables, early alcohol debut and educational achievements, as well as early cannabis initiation, will be addressed in the modeling strategy discussed in Section 4.

In addition to proneness, cultural, legal, economic, and physical accessibility can influence drug uptake. An indicator of the cultural accessibility of drugs could be the overall prevalence of illicit drug use in the society. Thus, prevalence data from another SIRUS study have been employed in the estimations, i.e data from annual surveys of 15-20 year old (corresponding surveys to the ones we use for young adults but conducted annually and for a younger target group). Since we have cross-sectional data, every person is assigned the prevalence rate that prevailed when he/she statistically was at the highest risk of taking up the various drugs. That means that every person is assigned the cannabis prevalence rate that applied when they were 18 years old, the amphetamine prevalence when they were 19, the heroin prevalence when they were 20 and the cocaine prevalence that applied when they were 21. Legal accessibility is assumed to be the same for all subjects since the same drug laws apply across the Norwegian counties. Drug prices can be used as indicators of economic accessibility. Unfortunately, we only have time series of drug prices for Oslo. Since regional police officers (personal communication) and field studies (Snertingdal 2007) have confirmed substantial price variations throughout the country, we do not use these price variables.

As for other indicators of accessibility, Table A1 also presents the responses to whether the young adults have ever been offered and can obtain (within 3 days) a range of different illegal drugs. It is interesting to note that half of the non-users of any illegal drugs report to have been offered cannabis and that almost 1 in 5 say they have been offered a hard drug. Also, roughly half of the cannabis-only users have been offered a hard drug (seemingly without accepting the offer). For the soft and hard drug users, however, the soft and hard drug offer variables are almost perfect predictors and the variables will therefore not be included in the statistical analysis in section 4. The various drug obtain variables,

displayed at the bottom of Table A1, are interesting as they can be interpreted in different ways. Firstly, they may give an indication of the physical availability of drugs. Not everyone claims to be able to get hold of narcotics within a relatively short time period, but roughly half of the non-users state they would be able to obtain cannabis and 4 in 10 say they could obtain a hard drug within 3 days. As expected, the corresponding numbers for illicit drug users were much higher. Secondly, regardless of actual drug use, knowing how to obtain an illicit commodity may provide additional information about the person (his/her specific knowledge, peer group, etc.) and thus the variable could be treated as an additional proneness variable. Since this information is given at the time of the survey, however, the variables may not be strictly exogenous. Our empirical analysis will therefore include models with and without the"obtain hard drug" ("hardobt") variable.

4 Data Analysis Methods

Our objective is to explore the role of prior cannabis use, proneness and accessibility on hard drug uptake in the Norwegian data, while accounting for unobserved proneness and accessibility and taking into account the strict frequential nature of the data. The data suggest that subjects first take the decision whether or not to enter the illicit drug market via cannabis which in turn affects their hard drug uptake. While cannabis users then decide whether or not to initiate hard drug use non-users stay away from illicit drugs. Further, given the the incomplete information about for proneness and accessibility in the data we can of course not rule out the presence of possible unobserved factors, e.g. genetic and environmental factors, that affect both cannabis and hard drug intake (unobserved confounders). To address these data features we propose a "degenerate sample selection model" with the following key features:

- cannabis use modeled for all subjects, hard drug use only for cannabis users,
- cannabis and hard drug use are modeled jointly to account for unobserved confounders,
- proneness and accessibility are considered explicitly.

The model is formulated under the Bayesian paradigm and is estimated using Markov Chain Monte Carlo (MCMC) methods. This allows us to employ the Bayesian predictive analysis to estimate the probability of hard drug uptake under cannabis use after accounting for observed and unobserved confounder and

to explore relative importance of the gateway, proneness and accessibility effects in explaining hard drug uptake among cannabis users.

4.1 Model for Cannabis and Hard Drug Use

In this section we formulate a model for the binary cannabis (cannabis) use $s_i = 0, 1$ and the binary hard drug use $h_i = 0, 1$ in the Norwegian Data, where one refers to drug use and zero to nonuse. We begin by modeling hard drug use for the cannabis users in the sample. Since a particular feature of illicit hard drug use in the (original) data is that cannabis use precedes hard drug use for almost all subjects in the sample, we assume that hard drug use can only occur if we observe cannabis use, $s_i = 1$, for a subject. If $s_i = 0$, then hard drug intake is fixed at $h_i = 0$. Thus, we have

$$h_i = \begin{cases} 0 & \text{if } s_i = 0 \\ (0, 1) & \text{if } s_i = 1 \end{cases}$$

To model the binary hard drug use for those individuals who are cannabis users, we introduce the latent continuous hard drug use variable h_i^* that is related to the observed binary hard drug use through the identity function as $h_i = I[h_i^* > 0]$. We specify the hard drug use in terms of the latent variable as

$$h_i^* = \mathbf{w}_{h,i}'\boldsymbol{\beta} + \varepsilon_i, \quad \varepsilon_i \sim N(0, 1), \tag{4.1}$$

where $\mathbf{w}_{h,i}$ is a $p \times 1$ vector of covariates. This implies, of course, a probit model for hard drug use with $\Pr(h_i = 1) = \Pr(h_i^* > 0) = \Phi(\mathbf{w}_{h,i}'\boldsymbol{\beta})$, where Φ is the cdf of a standard Normal distribution.

To evaluate the proneness and accessibility hypotheses we include, in addition to the usual demographic variables $(\mathbf{d}_{h,i})$, indicators to control for these effects in the covariate vector:

$$\mathbf{w}_{h,i} = (1, \mathbf{d}_{h,i}, \mathbf{p}_{h,i}, \mathbf{a}_{h,i}, g_{h,i}).$$

In particular, $\mathbf{p}_{h,i}$ will include indicators for problems with police, parents, family and friends and an indicator for starting with alcohol at an early age and dropping out of school early. Accessibility controls $\mathbf{a}_{h,i}$ will include the prevalence of the use of the various hard drugs and whether the respondent can obtain hard drugs. To capture the effect of early cannabis initiation, we also include $g_{h,i}$ to indicate early cannabis use (before age 16).

To account for the common problem that hard drug uptake among cannabis users may depend on unobserved factors that also drive cannabis use, we explicitly model cannabis use for all subjects in the

sample. In particular, we specify a probit model for the cannabis use s_i in terms of a corresponding latent continuous variable s_i^* as

$$s_i^* = \mathbf{w}'_{s,i}\boldsymbol{\gamma} + u_i, \quad u_i \sim N(0,1) \tag{4.2}$$

where $\mathbf{w}_{s,i}$ is a $q\times 1$ vector of observed characteristics, containing demographic, proneness and accessibility variables. For cannabis users, we then allow the error term to be correlated with that in the hard drug use equation. Specifically, we assume that (ε_i, u_i) are jointly normally distributed with mean zero and correlation matrix

$$\boldsymbol{\Omega} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

for subjects with $s_i = 1$. Here, the correlation parameter will correct for, among others, unobserved accessibility and proneness factors, such as genetic factors, that are not captured by the limited accessibility and proneness controls that are included in the covariate vectors.

From the above assumptions it follows directly that the joint model for the latent hard drug and cannabis use for cannabis users is the bivariate Normal distribution $\mathcal{N}_2(h_i^*, s_i^* | \mathbf{W}'_i\boldsymbol{\delta}, \boldsymbol{\Omega})$, where $\mathbf{W}_i = (\mathbf{w}_{h,i}, \mathbf{w}_{s,i})$ and $\boldsymbol{\delta} = (\boldsymbol{\beta}, \boldsymbol{\gamma})$. Thus, the likelihood contribution $\mathsf{Pr}(h_i = l, s_i = j | \mathbf{W}_i, \boldsymbol{\beta}, \gamma, \boldsymbol{\Omega})$ of the ith subject with $s_i = 1$, can be written in terms of the appropriate distribution of the latent drug variables (h_i^*, s_i^*) as

$$\mathsf{Pr}(h_i = l, s_i = 1 | \mathbf{W}_i, \boldsymbol{\beta}, \gamma, \boldsymbol{\Omega}) = \int_{A_l} \int_{A_1} \mathcal{N}_2(h_i^*, s_i^* | \mathbf{W}'_i\boldsymbol{\delta}, \boldsymbol{\Omega})\, ds_i^*\, dh_i^*$$

where $l = 0, 1$, and the integration regions are $A_0 = \{-\infty, 0\}$ and $A_1 = \{0, +\infty\}$. For subjects with $s_i = 0$, we only need to model the cannabis use, which can be expressed as the integral

$$\mathsf{Pr}(s_i = 0 | \mathbf{W}_i, \boldsymbol{\beta}, \gamma, \boldsymbol{\Omega}) = \int_{A_0} \mathcal{N}(s_i^* | \mathbf{w}'_{s,i}\boldsymbol{\gamma}, \mathbf{1})\, ds_i^*.$$

In anticipation of the discussion on the fitting of the model in the next section, we note that the likelihood contributions of the observed and the latent drug outcomes have a more attractive form. In particular, this joint likelihood contribution for subjects in with $s_i = 1$, $p(h_i^*, s_i^*, h_i = l, s_i = 1 | \mathbf{W}_i, \boldsymbol{\beta}, \gamma, \boldsymbol{\Omega})$, is

$$\mathcal{N}_2(h_i^*, s_i^* | \mathbf{W}'_i\,\boldsymbol{\delta}, \boldsymbol{\Omega})\, [I\{s_i^* \le 0\}^{1-s_i} + I\{s_i^* > 0\}^{s_i}]\, [I\{h_i^* \le 0\}^{1-h_i} + I\{h_i^* > 0\}^{h_i}] \tag{4.3}$$

while for subjects with $s_i = 0$ we have that $p(s_i^*, s_i = 0 | \mathbf{W}_i, \boldsymbol{\beta}, \gamma, \boldsymbol{\Omega})$ is of the form

$$\mathcal{N}(s_i^* | \mathbf{w}'_{s,i}\boldsymbol{\gamma}, \mathbf{1})\, [I\{s_i^* \le 0\}^{1-s_i} + I\{s_i^* > 0\}^{s_i}]. \tag{4.4}$$

Since we employ a Bayesian inferential approach we complete the model specification by defining the prior distribution for the vector of model parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)$. Following common practice, we specify normal prior distributions for the slope parameters. For the correlation coefficient we specify a normal prior that is restricted to the region $R = \{0 < \rho < 1\}$ to ensure the positive definiteness of the correlation matrix $\boldsymbol{\Omega}$ and to incorporate previously reported positive correlations between various drug intake variables (see for example Bretteville-Jensen et al, 2005). Assuming that the parameters are apriori independent, distribution of $\boldsymbol{\theta}$ is given by

$$\pi(\boldsymbol{\theta}) = \mathcal{N}_p(\boldsymbol{\beta}|\boldsymbol{b}_0, \mathbf{B}_0)\mathcal{N}_q(\boldsymbol{\gamma}|\boldsymbol{g}_0, \mathbf{G}_0)\mathcal{N}(\rho|r_o, R_0) \times R. \tag{4.5}$$

For the fitting of the model we center the prior distributions at zero for the slope parameters and at .2 for the correlation parameter. We further fix the prior variances at 5 for the slope coefficients and at 1 for the correlation parameter.

4.2 Model Fitting

In the Bayesian estimation framework all information about the model parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho)$ is summarized in the posterior distribution of the model parameters conditional on the observed data $\pi(\boldsymbol{\theta}|\mathbf{h}, \mathbf{s}, \mathbf{W})$. From Bayes theorem it follows that posterior distribution is proportional to the product of the prior distribution $\pi(\boldsymbol{\theta})$ and the likelihood function $p(\mathbf{h}, \mathbf{s}|\mathbf{W}, \boldsymbol{\theta})$. The goal is to develop a Markov Chain Monte Carlo algorithm to generate draws from the high dimensional posterior distribution (Chib, 2001). After a burn-in phase, the draws from the algorithm will converge to that of the posterior distribution. We can then summarize the posterior distribution of the model parameters based on these draws. Usually this is done in terms of the (posterior) means and standard deviations.

Due to the structure of the likelihood function, the posterior distribution for the problem at hand is not easily estimable via MCMC methods. We can however, exploit the ideas in Albert and Chib (1993) and include the latent drug intake variables in the parameter space and work with the posterior distribution $\pi(\boldsymbol{\theta}, \mathbf{s_i}^*, \mathbf{h_i}^*|\mathbf{h}, \mathbf{s}, \mathbf{W})$ which is given by the expression

$$\pi(\boldsymbol{\theta}, \mathbf{h_i}^*, \mathbf{s_i}^*|\mathbf{h}, \mathbf{s}, \mathbf{W}) \propto \pi(\boldsymbol{\theta})p(\mathbf{h}, \mathbf{s}, \mathbf{h_i}^*, \mathbf{s_i}^*|\mathbf{W}, \boldsymbol{\theta}),$$

where the last expression on the RHS is the joint likelihood for the observed and latent drug use outcomes. Assuming that the outcomes across individuals are distributed independently, it follows from expressions

(4.3) and (4.4) that this joint likelihood $p(\mathbf{h}, \mathbf{s}, \mathbf{h_i}^*, \mathbf{s_i}^* | \mathbf{W}, \boldsymbol{\theta})$ takes the form

$$\prod_{i \in N_0} \mathcal{N}(s_i^* | \mathbf{w}'_{s,i} \boldsymbol{\gamma}, \mathbf{1}) \left[I\{s_i^* \leq 0\}^{1-s_i} + I\{s_i^* > 0\}^{s_i} \right]$$
$$\prod_{i \in N_1} \mathcal{N}_2(h_i^*, s_i^* | \mathbf{W}'_i \boldsymbol{\delta}, \boldsymbol{\Omega}) \left[I\{s_i^* \leq 0\}^{1-s_i} + I\{s_i^* > 0\}^{s_i} \right] \left[I\{h_i^* \leq 0\}^{1-h_i} + I\{h_i^* > 0\}^{h_i} \right] \quad (4.6)$$

where N_j denotes the set of subjects using cannabis $j = 0, 1$, ie., $N_j = \{i : x_i = j\}$. The likelihood function is composed of two distinct terms, one from the $s_i = 0$ observations and the other from the $s_i = 1$ observations. The resulting joint posterior distribution is of a type that can be efficiently processed by Markov Chain Monte Carlo (MCMC) methods. We propose a six step MCMC chain to generate draws from the posterior distribution. The algorithm is provided in the appendix.

4.3 Predicting Hard Drug Uptake for Cannabis Users

An attractive feature of the Bayesian approach is that provides us with a tool for the predictive analysis, allowing us to assess the role of cannabis use, accessibility and proneness on the marginal probability of hard drug uptake. Given the structure of the data and our modeling strategy, the probability of hard drug uptake among cannabis non-users is zero. The key probability for our analysis is, therefore, the probability of hard drug uptake among cannabis users. Under the Bayesian predictive approach, we can compute the marginal predictive probability of hard drug uptake of $\mathsf{Pr}(h_{1,n+1} = 1 | \mathbf{h}, \mathbf{s}, \mathbf{W})$ of a random subject $(n+1)$ from the population. The probability only depends on the observed sample data $(\mathbf{h}, \mathbf{s}, \mathbf{W})$. The subscript 1 indicates that the probability is computed under the assumption of previous cannabis use. We can express this probability in terms of the latent hard drug variables $h^*_{1,n+1,}$ given the observed data as

$$\mathsf{Pr}(h_{1,n+1} = 1 | \mathbf{h}, \mathbf{s}, \mathbf{W}) = \mathsf{Pr}(h^*_{1,n+1} > 0 | \mathbf{h}, \mathbf{s}, \mathbf{W}) = \int_{A_1} p(h^*_{1,n+1} | \mathbf{h}, \mathbf{s}, \mathbf{W}) dh^*_{1,n+1} \quad (4.7)$$

where the expression under the integral is the marginal predictive distribution of the latent hard drug intake (for cannabis users). Note that the predictive distribution only depends on the observed data. While this expression is not immediately available but can be estimated by MCMC methods using the following decomposition

$$p(h^*_{1,n+1} | \mathbf{h}, \mathbf{s}, \mathbf{W}) = \int p(h^*_{1,n+1} | \mathbf{h}, \mathbf{s}, \mathbf{W}, \mathbf{w}_{h,n+1}, \boldsymbol{\beta}) \pi(\boldsymbol{\beta} | \mathbf{h}, \mathbf{s}, \mathbf{W}) p(\mathbf{w}_{h,n+1} | \mathbf{h}, \mathbf{s}, \mathbf{W}) d\boldsymbol{\beta} d\mathbf{w}_{h,n+1} \quad (4.8)$$

that expresses the marginal predictive distribution of latent hard drug use in terms of the distribution of latent hard drug use conditional on the parameters and the covariates which are marginalized out using

the posterior distribution and the empirical distribution, respectively. Based on this expression we can develop MCMC methods to generate M draws from the marginal distribution $(h_{1,n+1}^{*(1)}, h_{1,n+1}^{*(2)}, ..., h_{1,n+1}^{*(M)})$, which we then use to estimate the predictive probability of hard drug uptake for a cannabis user as follows:

$$\mathsf{Pr}(h_{1,n+1} = 1|\mathbf{h}, \mathbf{s}, \mathbf{W}) = \frac{1}{M}\sum_{g=1}^{M} I\{h_{1,n+1}^{*(g)} > 0\}. \tag{4.9}$$

Once we have obtained this quantity we can decompose the probability in terms of the gateway, proneness and accessibility effects as discussed in Section 4.3.

5 Results

We now present the analysis of the Norwegian Data by employing the methods just described. The binary dependent variable for cannabis use is set equal to 1 if the respondents report to have used cannabis at least 5 times. Many youngsters seem to try cannabis once or only a couple of times and we want to exclude those "accidental" users. For the hard drug use we consider two specifications. First, we again let the dependent variable equal 1 if subjects have used a hard drug at least five times, referred to as the Freq55 model hereafter. Second, since using a hard drug even once increases the risk of adverse health effects (e.g. the risk of overdose death for heroin use) and since a fraction of the sample is fairly young and may not yet have developed a more frequent drug using habit (the mean starting age for hard drug is higher than for cannabis), we also run a version of the model in which $h_i = 1$ for hard drug use ≥ 1 (Freq51 model). Further, we also employ two versions of the data set. One is the full data set described above and the second is the restricted version. As mentioned in Section 3, there is a group of hard drug users who stated the same starting age for cannabis and the hard drug. Thus, even though most people seem to start with cannabis, for these subjects we cannot strictly know which drug use came first. In the restricted version we therefore exclude those respondents, reducing the number of observations from 10,844 to 10,617.

5.1 Fitting of the Model

Table A2 shows the estimation results in terms of the posterior means and standard deviations of the parameters in the cannabis and hard drug equations. As mentioned in Section 3, we have estimated two versions of the model - one which includes the "hard drug obtain" variable among the controls for hard

drug uptake and one which does not include this variable. The results stem from the Freq55 model. Focusing on the most extensive version (the second result column), we see that the probability of having used a hard drug at least 5 times are higher among males, older age groups and people with low education. More surprisingly, living in the capital does not seem to have an effect on hard drug use after accounting for the other covariates. Turning to the proneness variables, problems with parents and police seem to increase the risk of using hard drugs at least five times, as does initiating alcohol use at a very young age. Table A1 showed that more than half of the cannabis users reported they could obtain hard drugs within 3 days and we find a large positive effect on hard drug use. The prevalence variables, on the other hand, are not as precisely estimated and the effects are small. The correlation term is positive indicating the presence of unobserved factors that affect both soft and hard drug use. This result is in line with the literature and supports our joint modeling strategy.

The last column in Table A2 presents the results from the less extensive model. Excluding the "hard drug obtain" variable does not change the signs of the estimated coefficients, but leads to small changes in some of the estimates. The main change occurs in the intercept which increases from -2.131 to -1.394 when going from the extensive model to the less extensive one. Table A2 shows that for the other coefficients, the absolute values for gender and year of survey increases to some extent, while the results for the other covariates stay roughly the same. The correlation term remains unchanged across the two models.

Table A2 further displays the results for the cannabis equation (similar for both models), and it is interesting to notice that gender, year of survey and living in Oslo all seem to have a larger impact on cannabis uptake than on hard drug initiation. Also the effect of the proneness variables differ between the two as problems with parents and school are more important for cannabis uptake whereas the police coefficient is larger for the hard drug models. To leave school early does in addition seem to affect the hard drug uptake more than it influences cannabis use. Early alcohol debut, on the other hand, has a greater impact on cannabis than on hard drug initiation.

(Table A2 about here)

5.2 Predicting the Probability of Hard Drug Uptake

In Table A3 we present the estimates for the predictive probability of hard drug uptake under the two different definitions of the hard drug use variable, for the two different model specifications and the two different samples. The first two columns refer to the case when the sequential model is fit to the full data set using the full covariate vector $\mathbf{w}_{h,n+1}$ and then the reduced covariate vector without the variable "obtain hard drugs", respectively. The last two columns give the corresponding results for the reduced sample where we excluded subjects that started with soft and hard drugs in the same year. The estimates were computed according to expression (4.9), based on draws of the (extended) MCMC algorithm that was run for 10,000 iterations (following a burn-in period of 1000 iterations). We find that within each model, the results are robust to an inclusion of the "obtain hard drug" variable and to the sample used. In the Freq55 model, the estimates for the total predictive probabilities range between 22% and 25%. As one would expect from the preliminary data analysis, the probabilities are higher under the corresponding Freq51 model, ranging between 37% and 40%.

(Table A3 about here)

These (total) predictive probabilities $\mathsf{Pr}(h_{1,n+1}|\mathbf{h}, \mathbf{s}, \mathbf{W})$ give the marginal probability of hard drug uptake when a random subject from the sample would be exposed to cannabis at least five times. They incorporate the uncertainty about the parameter values and the covariates in a statistical coherent way via marginalization using the corresponding posterior density of the parameters (from the model fitting) and the empirical distribution of the data (see expression 4.8). The probabilities itself have no variance. However, we could (and will in later version) provide density plots of the predictive distributions and report the means and variances for the underlying predictive distributions of the latent hard drug use variable $h^*_{i,n+1}$ for an additional comparison of the hard drug uptake across the different specifications.

5.3 Gateway, Proneness and Accessibility Effects

While the total predictive probabilities are interesting by themselves, we are also interested in a decomposition of the total probability in terms of the gateway, proneness and accessibility effects. One strength of our model is that it includes controls for proneness and accessibility, in addition to standard demographic characteristics. To quantify the effect of cannabis use on hard drug uptake, we start by predicting the

hard drug uptake under the modified covariate vector $\tilde{\mathbf{w}}_{h,n+1}$ that only contains the intercept. Since there is a possibility that cannabis use interacts with demographic characteristics we can repeat the analysis with $\tilde{\mathbf{w}}_{h,n+1} = (1, \mathbf{d}_{h,n+1})$. To include the transmission channel through early cannabis use, we let $\tilde{\mathbf{w}}_{h,n+1} = (1, g_{h,n+1})$ and $\tilde{\mathbf{w}}_{h,n+1} = (1, \mathbf{d}_{h,n+1}, g_{h,n+1})$, respectively.

We can exploit the same idea to assess the effects of proneness and accessibility on hard drug uptake. For example, to see whether proneness increases the probability of hard drug uptake, we let $\tilde{\mathbf{w}}_{h,n+1} = (1, \mathbf{d}_{h,n+1}, \mathbf{p}_{h,n+1}, g_{h,n+1})$. To check whether and by how much accessibility to drugs increases hard drug use we predict hard drug uptake under $\tilde{\mathbf{w}}_{h,n+1} = (1, \mathbf{d}_{h,n+1}, \mathbf{a}_{h,n+1}, g_{h,n+1})$. Table 1 gives the different

		Variables included in $\tilde{\mathbf{w}}_{h,n+1}$				
Hypothesis		intercept	$g_{h,n+1}$	$\mathbf{p}_{h,n+1}$	$\mathbf{a}_{h,n+1}$	$\mathbf{d}_{h,n+1}$
Gateway	G 1	x	-	-	-	-
	G 2	x	-	-	-	x
	G 3	x	x	-	-	-
	G 4	x	x	-	-	x
Gateway and Proneness	GP	x	x	x	-	x
Gateway and Accessibility	GA	x	x	-	x	-
Total (GPA)	GPA	x	x	x	x	x

Table 1: *Specifications of the covariate vector for the prediction of hard drug uptake to assess the three hypotheses.*

specifications of the covariate vector used in the prediction of the probability for hard drug uptake we use to assess the gateway, proneness and accessibility effects in the Norwegian data. The last expression in the table (GPA), corresponds to the estimate of the total predictive probability of hard drug uptake.

Estimates for the predictive probabilities in Table 1 for the full sample are provided in Table A4. The first two columns give the results for the Freq55 model when the variable "hardobt" is included in the analysis and when it is excluded. The next two columns give the corresponding estimates for the Freq51 model. If we only consider the effect of the intercept (G1), we find a probability of hard drug uptake between 2% and 7% under the Freq55 and Freq51 models (including the hardobt variable in estimation set-up), respectively. When this variable is omitted, then the corresponding effects increase to 8% and 16%. In this case the positive effect from the variable is (partially) captured by an increase of the intercept estimated as shown in Table A2.

The effect of early cannabis (G3) use ranges between 4.3% and almost 20%. For the Freq55 model, the combined effect (G4) ranges between 4.4% when the obtain variable is included in the model and 15% when it is excluded. Adding the effect of demographic characteristics increases the probability but the effect is small in relation to G1 and G3. These estimates suggest a positive gateway effect. While we cannot isolate the gateway effect directly, the estimate of G4 can be interpreted as an upper bound to a restrictively defined gateway effect in terms of a constant hurdle effect from cannabis use across subjects (intercept) and a heterogenous component from early cannabis initiation and a possible an interaction with personal characteristics. The last component is the smallest.

(Table A4 about here)

We now turn to the proneness and accessibility effects. If we include the proneness measured in terms of the various childhood problems and early alcohol use, the probability of hard drug use increases across all specification. The difference between GP1 and G4 implies proneness effects between 1.3% and 5.8%. As mentioned in Section 3.2, the variable for obtaining hard drugs can be interpreted as an indicator for accessibility and proneness. If we include it in the set of proneness variables, the probabilities increase by around 14%. To evaluate the increase in probability from accessibility defined in terms of the drug prevalence, we turn to the next estimate in the table (GA). We find that including effects from the prevalence variables increases the probability of hard drug uptake between 1.3% and 10.3% compared to G4. By including the hardobt variable in the set of proneness indicators, the accessibility effect is relatively small. There is of course the possibility that "hardobt" captures the interaction of both effects. To summarize, our decomposition indicates the presence of all three effects. The relative importance of the effects partially depends on whether the variable to obtain hard drugs is included in the analysis, and what effect it presents (proneness, accessibility or a combination of both). Note again, that the hardobt variable does not influence the estimates of the total predictive probabilities. To investigate the robustness of these results we have repeated the analysis from Table A4 under the restricted sample. Our general findings are robust with respect to the sample choice. The results are available upon request.

5.4 Policy Experiments

In the previous section we computed the predictive probabilities under different specifications of the covariate vector $\tilde{\mathbf{w}}_{h,n+1}$ to explore the gateway, proneness and accessibility effects on the hard drug uptake. In each case we used the empirical distribution of the data to marginalize over the included covariates. We can, however, also use our estimation set-up to explore the hard drug uptake under different policy relevant scenarios, see Table A5.

(Table A5 about here)

For instance, we can compute the overall probability of hard drug uptake under the assumption that it is possible to prevent early initiation of cannabis use, i.e. setting the "start young" variable equal to 0 for all subjects while keeping the sample values for all other variables constant. Results from this hypothetical exercise could be interpreted as an upper bound of what might happen if drug policy aiming at increasing the debut age of cannabis use were highly successful. Table A5 shows a decrease in overall probability of about 4 percentage points across the different models and sample sizes. If, on the other hand, every cannabis user initiated his or her use before turning 16 years (the "start young" variable set equal to 1 for all subjects), the probability of hard drug uptake would increase to roughly 40 per cent for the Freq55 model and to 53 per cent for the Freq51 model. More people starting at an early age might be a likely consequence of cannabis legalization if cannabis then is used in a similar way as alcohol (recall that the average debut age for alcohol in the current sample is 15.4 versus 18.6 for cannabis). Table A5 also shows the changes in overall probabilities from setting all the problem variables to 0, setting one at a time to 1 and setting all to 1, respectively. In line with the results presented above, Table A5 suggests that a change in the police variable again has the largest impact on hard drug uptake. If every young adult had reported childhood problems with parents, school, police and friends the probability of initiating hard drugs would, according to these results, would be 50 per cent for the Freq55 model. The last two rows of Table A5 present overall probabilities given that the prevalence variables were given low values (0) and high values (prevalence for amphetamine set equal to 25 per cent, cocaine to 20 per cent and heroin to 10 per cent, respectively).

5.5 Additional Results and Sensitivity Analysis

To be added: brief discussion of results from the fitting of the Freq55 model to the Oslo data, the full sample restricted to subjects that are over 25 years old (Table A6) and the fitting of the Freq51 model for amphetamine, cocaine and ecstacy use (Table A7).

(Table A6 and Table A7 about here)

6 Discussion

Hard drug use inflicts considerable harm on drug users and society. Therefore, a wide range of policy means have been adopted to confine drug consumption and prevent people from taking up the habit. Motivated by the staircase pattern of drug use, previous research and policy discussions have focused on the gateway hypothesis that cannabis use directly increases the risk of hard drug use. Many countries have implemented strict cannabis laws based on the gateway hypothesis. However, the staircase pattern can also be explained by two alternative effects that have been discussed in the literature, the proneness and the accessibility effects. This paper uses a unique Norwegian data set and a Bayesian predictive framework to explore the relative importance of the three effects in the hard drug uptake among cannabis users. Given the frequential nature of the data, we propose a degenerate sample selection model that accounts for unobserved confounders as basis for the predictive analysis.

The empirical analysis in this paper suggests a direct effect of cannabis use on hard drug initiation, after accounting for observed and unobserved proneness and accessibility, across various model specifications and sample definitions. Thus, our results are in line with DeSimone (1998), van Ours (2003), Bretteville-Jensen et al. (2005), Furgusson et al. (2006) and Melberg et al. (2007), and contrast the findings of Benstock and Rahav (2002) and Pudney (2003). The results in Bretteville-Jensen et al. (2005) and Melberg et al. (2007) are based on the Oslo sample of the 2002 survey. Further, we find that proneness and accessibility factors matter in explaining the hard drug uptake among cannabis users. The magnitude of the three effects, but less so the relative importance of the hypotheses, vary across model specifications. We assume that the most policy relevant model is the one in which the dependent variable is set equal to one if the respondent has used a drug 5 times or more. Results from this model specification (Freq55) suggest that a cannabis user has a probability of initiating hard drug use of roughly 14 per cent. De-

composing this predicted probability in terms of the gateway, proneness and accessibility effects suggest that the gateway effect raises the probability by 2.2 %, while adding proneness and accessibility effects increases the probability further to 2.9% and 4.1 % (14.1 % including hard obtain variable), respectively. Table A4 showed that the "hardobt" variable made a significant difference to the relative importance of the three explanatory categories but that the estimate for overall probability was robust to the inclusion of the variable.

The paper also examines hypothetical policy scenarios to explore how various policies that affect for example the starting age of drugs, accessibility and proneness based on our estimation results. The preliminary results suggest, for example, that a successful intervention of rising the average age of cannabis initiation could considerably reduce the overall probability of hard drug uptake. A policy aimed at reducing childhood problems does not seem to be as effective as one decreasing the accessibility of drugs in limiting hard drug abuse. While proneness factors as these have a large effect on a subject's probability of hard drug uptake there is a low prevalence of these factors in the data. It is likely that our results underestimate policy effects that is targeted to problem groups as our current analysis is based on the full sample. Policy implications of the gateway effect are less clear. Despite what is commonly assumed, finding a gateway effect does not necessarily imply that a strict cannabis policy would be the preferred option. The policy response will depend on which mechanisms are behind the gateway effect. So far, four such mechanisms have been suggested: (1) the drug use itself induces a craving for more and stronger drug experiences (altered preferences), (2) contact with hard drug users and dealers could lead cannabis users to proceed to the next "step/level" (social interaction), (3) experiencing no/little adverse effects of cannabis use also reduces the confidence in health warnings against hard drugs (increased information), or (4) the "costs" of using an illegal drug is substantially reduced after first having broken the law and social norms by using the illegal cannabis (reduced costs). If altered preferences, increased information or reduced costs are the driving forces then a strict cannabis policy may seem appropriate, whereas if contact with hard drug dealers increases the probability of hard drug uptake, then separating the markets for soft and hard drugs, in line with the Dutch drug policy, may be preferred. Unfortunately, the data at hand are not rich enough to pin down which mechanism(s) that is (are) at work here.

The paper has several shortcomings. The survey does not provide us with all the variables of interest,

e.g. it would have been useful with more information regarding an individuals' pre-drug-using period. As a result, it can be argued that some of the variables may be endogenous. Acknowledging that no survey ever will include data on every possible factor that could influence drug initiation, we have, like the recent literature on this topic, adopted an approach that accounts for unobserved heterogeneity. The approach implies that the unobserved are assumed to be individual specific and stable over time. However, if these assumption are violated, the econometric method may not fully control for the unobserved factors, however. Further, the relatively low response rate and the possible under-representation of problematic drug users may limit the generalizability of our findings. The sample may be more representative for "most young adults" than "all young adults". The results indicate, however, that problems during childhood etc. increases the probability of hard drug initiation. Given the higher problem prevalence figures for drug treatment populations (Lauritzen et al. 2003), one would expect that our findings would be strengthened if more problematic drug users were too be included in the sample.

References

Albert, J. and Chib, S. Bayesian Analysis of Binary and Polychotomous Response Data 1993; 88; 669-779.

Beenstock M, Rahav, G. Testing Gateway Theory: do cigarette prices affect illicit drug use? Journal of Health Economics 2002; 21: 679-698.

Becker G.S, Murphy K.M. A Theory of Rational Addiction. Journal of Political Economy 1988; 96: 675-700.

Bretteville-Jensen A.L, Melberg H.O, Jones A. Sequential patterns of drug use initiation - can we believe in the gateway theory? HEDG working paper no 05/09, Department of Economics and Related Studies, University of York, UK, 2005.

DeSimone J. Is marijuana a gateway drug? Eastern Economic Journal 1998; 24: 149-164.

Ellgren M, Spano, S.M, Hurd Y.L. Adolescent cannabis exposure alters opiate intake and opioid limbic neuronal populations in adult rats. Neuropsychopharmacology 2007; 32: 607-615.

EMCDDA. Annual report 2006: the state of the drugs problems in Europe. European Monitoring Centre for Drugs use and Drug Addiction, Lisbon, Portugal, 2006.

Fergusson DM, Boden JM and Horwood LJ. Cannabis use and other illicit drug use: testing the cannabis gateway hypothesis. Addiction 2006; 101: 556-569.

Kandel D.B. Stages in adolescent involvement in drug use. Science 1975; 190: 912-914.

Kandel D.B (ed.). Stages and pathways of drug involvement: Examining the gateway hypothesis. Cambridge, England: Cambridge University Press, 2002.

Kenkel D, Mathios A.D, Pacula R.L. Economics of youth drug use, addiction and gateway effects. Addiction 2001; 96: 151-164.

Lauritzen G, Waal H, Amundsen A and Arner O. A national study of Norwegian drug abusers in treatment. Methods and findings. Nordic Studies on Alcohol and Drugs (NAT) 1997; 14: 43-63.

Lessem J.M, Hopfer C.J, Haberstick B.C, Timberlake D, Ehringer M.A, Smolen A, Hewitt J.K. Relationship between adolescent marijuana use and young adult illicit drug use. Behavior Genetics 2006; 36(4): 498-506.

Lynskey M.T, Heath A.C, Bucholz K.K, Slutske W.S, Madden P.A, Nelson E.C. Escalation of drug use in early-onset cannabis use versus co-twin controls. JAMA 2003; 289: 427-433.

Lynskey M.T, Vink J.M, Boomsma D.I. Early onset of cannabis use and progression to other drug use in a sample of Deutch twins. Behavior Genetics 2006; 36 (2): 195-199.

Melberg H.O, Bretteville-Jensen A.L, Jones A. Is cannabis a gateway to hard drugs? HEDG working paper no 07/01, Department of Economics and Related Studies, University of York, UK, 2007.

Morral A.R, McCaffrey D.F, Paddock S.M. Reassessing the marijuana gateway effect. Addiction 2002; 97: 1493-1504.

Nutt D, King L.A, Saulsbury W, Blakemore C. Development of a rational scale to assess the harm of drugs of potential misuse. The Lancet 2007; 369 (9566): 1047-1053.

Ours van J. Is cannabis a stepping-stone for cocaine? Journal of Health Economics 2003; 22: 539-554.

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

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

Skog O.J. The validity of self-reported drug use. British Journal of Addiction 1992; 87: 539-548.

Snertingdal M.I. Kalkulerende kjeltringer eller offer for omstendighetene: en kvalitativ studie av heroinomsetningens utvikling og aktører (Crocs or victims of circumstances: Qualitative study of the development of the Norwegian heroin market and its actors), SIRUS Report no.1/07, Norwegian Institute for Alcohol and Drug Research, Oslo, 2007.

Tsuang M.T, Lyons M.J, Meyer J.M, Doyle T, Eisen S.A, Goldberg J., et al. Co-occurrence of abuse of different drugs in men. Arch Gen Psychiatry 1998; 55: 967-972.

7 Appendix: Markov Chain Monte Carlo Methods

We generate draws from the posterior distribution of interest based on the following MCMC steps that are repeated 10,000 times following a burn-in period of 1000 draws. We let $\mathbf{s} = (\mathbf{s}_0, \mathbf{s}_1)$ where $\mathbf{s}_0 = (s_i : i \in N_0)$ and $\mathbf{s}_1 = (s_i : i \in N_1)$ represent the cannabis observations under the two possible cannabis states; with a similar convention for the latent cannabis intake is $\mathbf{s}^* = (\mathbf{s}_0^*, \mathbf{s}_1^*)$.

MCMC Algorithm 1

1. Initialize $\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho, \{s_i^*\}$
2. Sample $h_i^*|h_i, s_i, s_i^*, \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho$ for all subjects in N_1 from the conditional truncated normal $N(h_i^*|\mathbf{w}'_{h,i}\boldsymbol{\beta} + \boldsymbol{\rho}(\boldsymbol{s}_i^* - \mathbf{w}'_{s,i}\boldsymbol{\gamma}), 1-\rho^2)\,[I\{y_i^* \leq 0\}^{1-y_i} + I\{y_i^* > 0\}^{y_i}]$
3. Sample $s_i^*|h_i, s_i, h_i^*, \boldsymbol{\beta}, \boldsymbol{\gamma}, \rho$ from
   (a) $N(s_i^*|\mathbf{w}'_{s,i}\boldsymbol{\gamma}, 1)[I\{s_i^* \leq 0\}$ for all subjects with $s_i = 0$
   (b) $N(s_i^*|\mathbf{w}'_{s,i}\boldsymbol{\gamma} + \boldsymbol{\rho}(\boldsymbol{h}_i^* - \mathbf{w}'_{h,i}\boldsymbol{\beta}), 1-\rho^2)[I\{s_i^* > 0\}$ for all subjects with $s_i = 1$
4. Sample $\boldsymbol{\beta}|\mathbf{h}, \mathbf{s}_1, \mathbf{h}^*, \mathbf{s}_1^*, \boldsymbol{\gamma}, \rho$ from $\mathcal{N}_p(\boldsymbol{\beta}|\hat{\boldsymbol{\beta}}, \mathbf{B})$, where
   $\hat{\boldsymbol{\beta}} = \mathbf{B}[\mathbf{B}_0^{-1}\mathbf{b}_0 + \sum_{i\in N_1}\mathbf{w}_{h,i}\sigma^{-2}(h_i^* - \rho\hat{s_i}^*)]$,
   $\mathbf{B} = [\mathbf{B}_0^{-1} + \sum_{i\in N_1}\mathbf{w}_{h,i}\sigma^{-2}\mathbf{w}'_{h,i}]^{-1}$ with $\sigma^2 = (1-\rho^2)$,
   and $\hat{s_i}^* = s_i^* - \mathbf{w}'_{s,i}\boldsymbol{\gamma}$
5. Sample $\boldsymbol{\gamma}|\mathbf{h}, \mathbf{s}, \mathbf{h}^*, \mathbf{s}^*, \boldsymbol{\beta}, \rho$ from $\mathcal{N}_k(\boldsymbol{\gamma}|\hat{\boldsymbol{\gamma}}, \mathbf{G})$, where
   $\hat{\boldsymbol{\gamma}} = \mathbf{G}[\mathbf{G}_0^{-1}\mathbf{g}_0 + \sum_{i\in N_0}\mathbf{w}_{s,i}s_i^* + \sum_{i\in N_1}\mathbf{w}_{s,i}\sigma^{-2}(s_i^* - \rho\hat{h_i}^*)]$,
   $\mathbf{G} = [\mathbf{G}_0^{-1} + \sum_{i\in N_0}\mathbf{w}_{s,i}\mathbf{w}'_{s,i} + \sum_{i\in N_1}\mathbf{w}_{s,i}\sigma^{-2}\mathbf{w}'_{s,i}]^{-1}$,
   with $\sigma^2 = (1-\rho^2)$ and $\hat{h_i}^* = h_i^* - \mathbf{w}'_{h,i}\boldsymbol{\beta}$
6. Sample $\rho|\mathbf{h}_1, \mathbf{s}_1, \mathbf{h}_1^*, \mathbf{s}_1^*, \boldsymbol{\beta}, \boldsymbol{\gamma}$, for $j = 0, 1$ from an MH step described below.
7. Goto 2

Since the full conditional distribution of the correlation coefficient ρ is not tractable, we update the parameters using the Metropolis Hastings algorithm. Following Chib and Greenberg (1995,1998) we generate proposal values ρ' from a tailored student-t density $t_\nu(\mu, V)$ where μ is the approximate mode of

$$\ln\Big(\prod_{I\in I_1}\mathcal{N}(h_i^*, s_i^*|\mathbf{W}_i\boldsymbol{\delta}, \boldsymbol{\Omega})$$

and V is the inverse Hessian of this density evaluated at μ. We accept the proposal value with probability of move α where

$$\alpha = \left\{1, \frac{\pi(\rho')\prod_{i\in I_1}\mathcal{N}(h_i^*, s_i^*|\mathbf{W}_i\boldsymbol{\delta}, \boldsymbol{\Omega'}) \times t_\nu(\rho|\mu, V)}{\pi(\rho)\prod_{i\in I_1}\mathcal{N}(h_i^*, s_i^*|\mathbf{W}_i\boldsymbol{\delta}, \lambda_i^{-1}\boldsymbol{\Omega}) \times t_\nu(\rho'|\mu, V)}\right\}$$

8 Appendix: Tables and Figures

Figure A1. Initiation pattern for drugs. Mean starting ages among 21-30 year olds in Norway. Data collected in 2002 and 2006

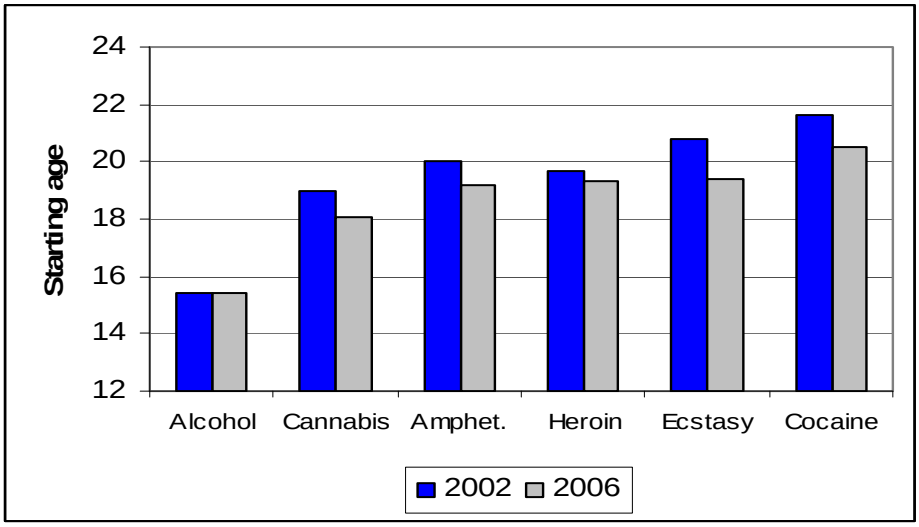

Table A1. Variable descriptions for the full sample and subgroups divided by extent of illegal drug use

	All (n=10,844)	No illegal drugs (n=6,957)	Illegal drugs (n=3,887)	Soft, not hard drugs (n=2,608)	Soft and hard drugs (n=1,279)
Cannabis use (1 if ever used)	0.36	0	1	1	1
Hard drug use (1 if ever used)	0.12	0	0.33	0	1
Cannabis young (1 if started to use cannabis ≤ 15)	0.06	0	0.16	0.09	0.30
Cannabis frequency (1 if used cannabis ≥ 25 times)	0.12	0	0.32	0.16	0.63
Gender (1 if male)	0.40	0.36	0.47	0.44	0.53
Age (in years)	26.4	26.5	26.4	26.4	26.2
Oslo (1 if living in Oslo)	0.62	0.56	0.71	0.71	0.72
Year (1 if surveyed in 2006)	0.38	0.37	0.40	0.40	0.40
Parents (1 if childhood problems)	0.08	0.06	0.13	0.09	0.19
School (1 if childhood problems)	0.06	0.04	0.10	0.07	0.16
Police (1 if childhood problems)	0.014	0.006	0.029	0.007	0.07
Friends (1 if childhood problems)	0.04	0.03	0.04	0.04	0.06
Problems (1 if any of the childhood problems)	0.13	0.10	0.20	0.15	0.30
Leave school early (1 if left school ≤ 16)	0.04	0.04	0.06	0.03	0.11
Alcohol young (1 if started to use alcohol ≤ 12)	0.06	0.04	0.11	0.07	0.18
Cannabis prevalence	22.2	22.1	22.3	22.3	22.5
Amphetamine prevalence	5.3	5.3	5.4	5.4	5.5
Cocaine prevalence	3.6	3.6	3.7	3.6	3.7
Heroin prevalence	1.2	1.2	1.2	1.2	1.2
Offered cannabis	0.67	0.50	0.99	0.99	0.99
Offered hard drugs	0.35	0.19	0.65	0.48	0.99
Can obtain cannabis	0.66	0.54	0.88	0.85	0.95
Can obtain hard drugs	0.48	0.39	0.63	0.52	0.85

Table A2. Results for the cannabis and hard drug equations. Dependent variables = 1 if the drug is used ≥ 5 times. Posterior means and standard deviations (in parentheses)

		Cannabis use	**Hard drug use** (including the hard drug obtain variable)	**Hard drug use** (excluding the hard drug obtain variable)
	Intercept	-1.739 (0.217)	-2.131 (0.336)	-1.394 (0.341)
Demographics:	**Gender**	0.279 (0.028)	0.119 (0.065)	0.192 (0.062)
	Age	0.001 (0.014)	0.049 (0.022)	0.041 (0.021)
	Age squard	-0.002 (0.002)	-0.003 (0.005)	-0.004 (0.005)
	Oslo	0.390 (0.030)	-0.020 (0.076)	-0.012 (0.076)
	Year	0.101 (0.041)	-0.067 (0.078)	-0.122 (0.077)
Proneness:	**Parents prob.**	0.489 (0.050)	0.235 (0.095)	0.262 (0.092)
	Police prob.	0.649 (0.114)	0.813 (0.168)	0.870 (0.163)
	Friends prob.	-0.177 (0.079)	-0.266 (0.151)	-0.276 (0.145)
	School prob.	0.422 (0.062)	0.138 (0.108)	0.169 (0.107)
	Leave school early	0.148 (0.066)	0.585 (0.116)	0.681 (0.115)
	Alcohol young	0.611 (0.052)	0.269 (0.099)	0.322 (0.096)
Accessibility variables:	**Cannabis prevalence**	0.018 (0.010)		
	Amphetamine prevalence		0.029 (0.034)	0.032 (0.032)
	Cocaine prevalence		0.030 (0.052)	0.046 (0.050)
	Heroin prevalence		-0.050 (0.111)	-0.043 (0.104)
	Obtain hard drugs		1.107 (0.090)	
	Cannabis young		0.650 (0.073)	0.704 (0.071)
Correlation term			0.180 (0.120)	0.184 (0.129)

Table A3. Predicted probabilities of hard drug uptake for Model55 and Model51 under different sets of covariates and samples

	Full sample (n=10,844)		**Restricted sample (n=10,617)**	
	Including the hard drug obtain variable	Excluding the hard drug obtain variable	Including the hard drug obtain variable	Excluding the hard drug obtain variable
Model Freq 55	0.138	0.167	0.122	0.145
Model Freq 51	0.267	0.305	0.248	0.280

Table A4. Decomposition of the predicted probabilities of hard drug uptake across the different hypotheses. (Full Sample)

		Freq55 (including hardobt variable)	**Freq55** (excluding hardobt variable)	**Freq51** (including hardobt variable)	**Freq51** (excluding hardobt variable)
Gateway	**G1**	0.0180	0.0870	0.0756	0.1589
	G2	0.2200	0.0951	0.0799	0.1688
+ Proneness	**GP1**	0.0290	0.1167	0.1004	0.2015
+ Accessibility	**GPA1**	0.0406	0.1670	0.1379	0.2838
(with hardobt)	**GPA2**	0.1406	-	0.2498	-
Total (Gateway)	**GPA**	0.1382	0.1667	0.2673	0.3047

Figure A2. Decomposition of distribution of latent predictive hard drug uptake across the different hypotheses for the Freq55 model.

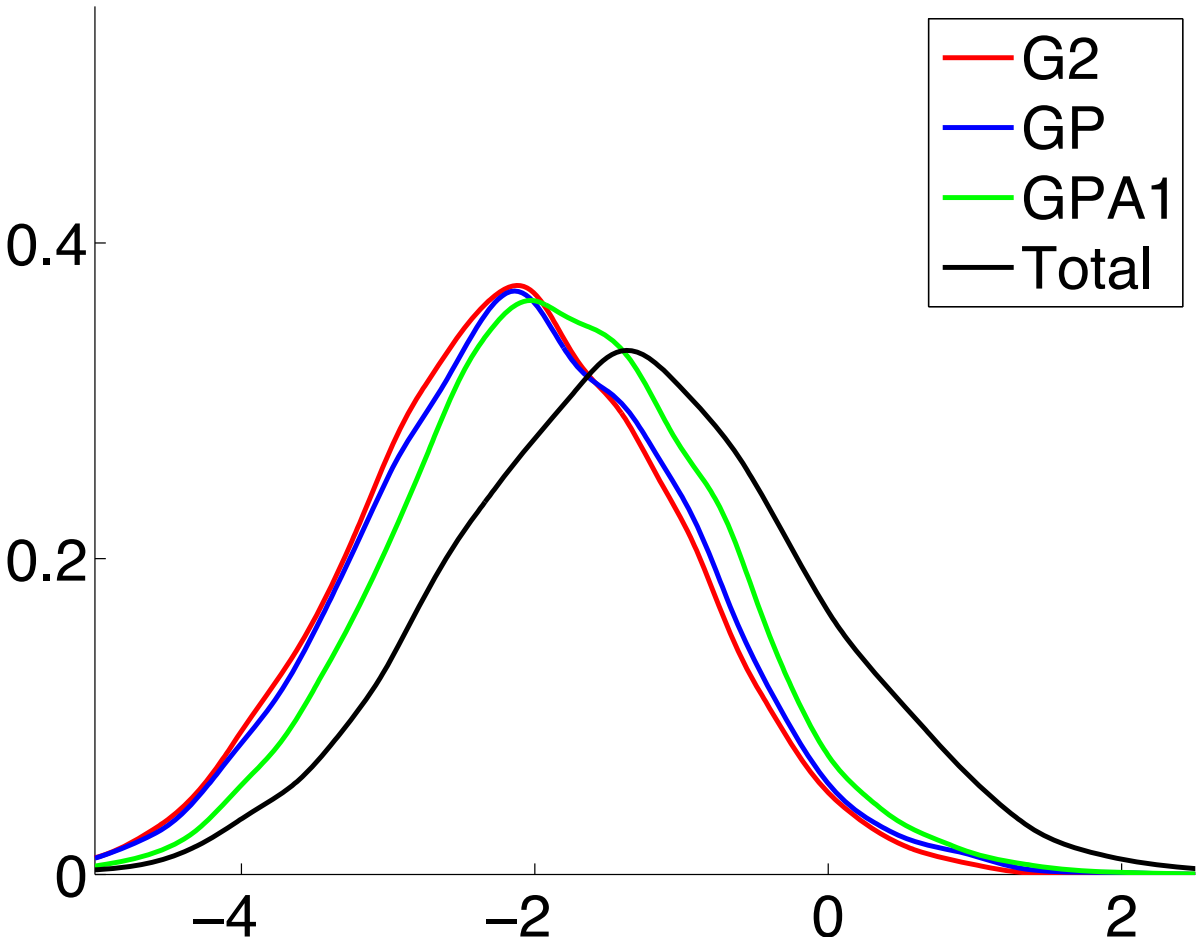

Figure A3. Decomposition of distribution of latent predictive hard drug uptake across the different hypotheses for the Freq51 model.

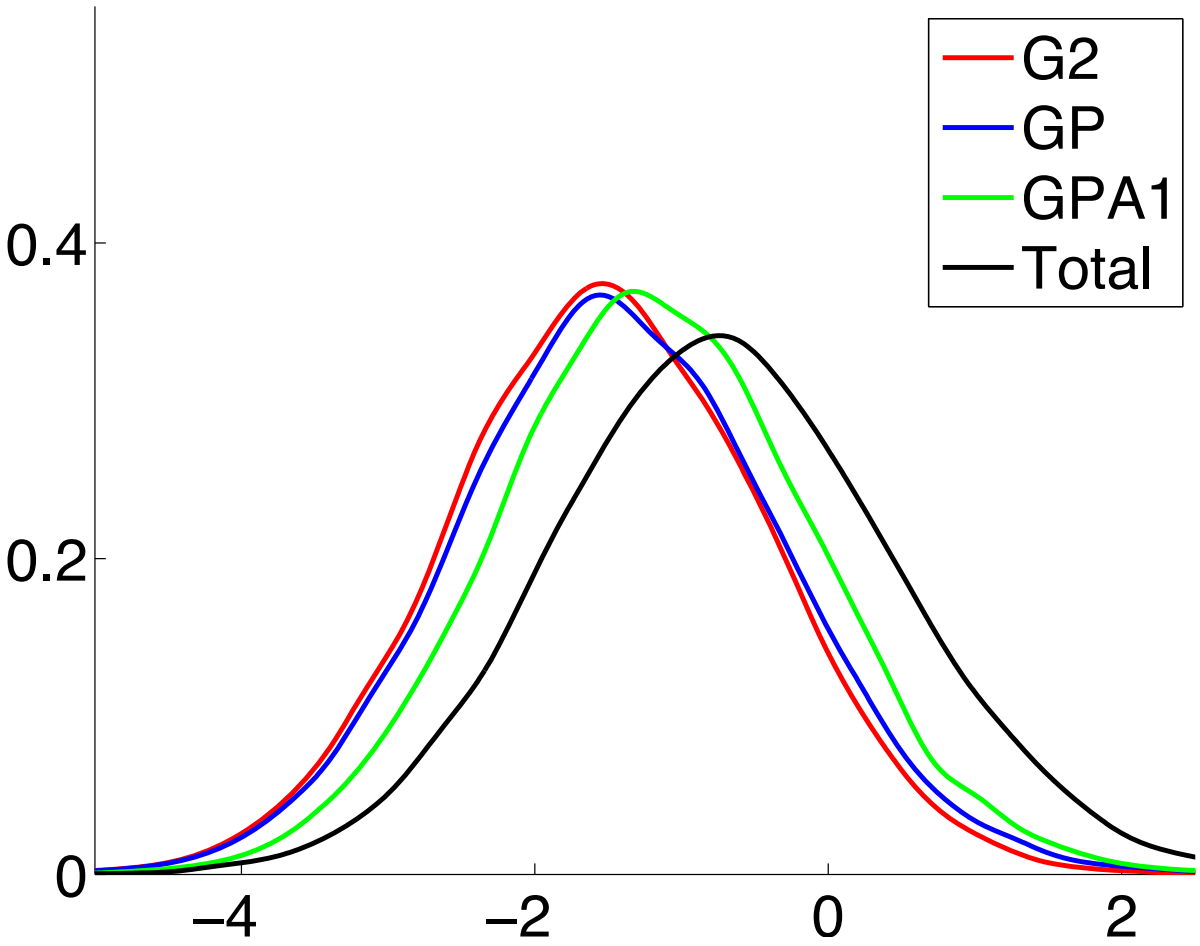

Table A5. The estimated predicted probabilities of taking up hard drug under different policy scenarios for the full (n=10,844) and the restricted sample (n=10,617)

	Freq55 Full sample	**Freq55** Restr. sample	**Freq51** Full sample	**Freq51** Restr. sample
Prob. of taking up hard drugs	0.138	*0.230*	0.267	*0.368*
Prob. of taking up hard drugs if				
Starting young = 0	0.217	*0.194*	0.261	*0.322*
Starting young = 1	0.270	*0.395*	0.419	*0.527*
Prob. of taking up hard drugs if				
Problems Police and Parents= 0	0.129	*0.211*	0.262	*0.349*
Police = 1, Parents = 0	0.324	*0.408*	0.625	*0.671*
Parents = 1, Police = 0	0.172	*0.262*	0.324	*0.416*
Prob. of taking up hard drugs if				
Leave school early = 0	0.131		0.263	
Alcohol young = 0	0.132		0.265	
Prob. of taking up hard drugs if				
Accessibility = Low	0.031	*0.056*	0.107	*0.148*
Accessibility = High	0.189	*0.232*	0.321	*0.374*

Table A6 (sensitivity analysis). Decomposition of the predicted probabilities for (1) subjects 26 years and above and (2) for subjects residing in Oslo

		(1) Freq55	(1) Freq51	(2) Freq55	(2) Freq51
Gateway	**G1**	0.0384	0.1435	0.0223	0.1010
	G2	0.0407	0.1481	0.0259	0.1091
+ Proneness	**GP**	0.0515	0.1715	0.0347	0.1091
+ Accessibility	**GPA1**	0.0559	0.1738	0.0482	0.1590
(including obt)	**GPA2**	0.1519	0.2829	0.1550	0.2793
Total (+ Demographics)	**GPA**	0.1410	0.2913	0.1560	0.2908

Table A7(sensitivity analysis). Decomposition of the predicted probabilities of amphetamine, cocaine and ecstasy uptake across the different hypotheses. (Full Sample)

		Freq51 Amphetamine	**Freq51** Cocaine	**Freq51** Ecstasy
Gateway	**G1**	0.0892	0.0862	0.0747
	G2	0.0930	0.0902	0.0783
+ Proneness	**GP**	0.1080	0.1047	0.0918
+ Accessibility	**GPA1**	0.1256	0.1316	0.0963
(including obt)	**GPA2**	0.2241	0.1986	0.1447
Total (+ Demographics)	**GPA**	0.1933	0.1791	0.1201