THE ROAD TO RUIN? SEQUENCES OF INITIATION TO DRUG USE AND OFFENDING BY YOUNG PEOPLE IN BRITAIN *

Stephen Pudney
Department of Economics, University of Leicester

April 2001, revised January 2002

ABSTRACT: The routes by which young people develop offending behaviour are very varied and strongly influenced by family background. A good understanding of the temporal sequences of first experiences of illicit drug use and other offending behaviour is needed before any plausible attempt can be made to investigate causal "gateway" effects. In this paper we develop and apply a statistical method for analysing the behavioural sequences observed in the 1998 Youth Lifestyles Survey. Gateway effects are found to be small after controlling for observable and unobservable characteristics.

KEYWORDS: illicit drugs, gateway effect, youth crime, random effects.

JEL CLASSIFICATION: I120, K420

ADDRESS FOR CORRESPONDENCE: Stephen Pudney, Department of Economics, University of Leicester, University Road, Leicester LE1 7RH, UK; e-mail: sep2@le.ac.uk.

* I am grateful to the Home Office for seedcorn funding for this research and for providing access to the Youth Lifestyles Survey data. All errors are mine, as are the views expressed here.

1 Introduction

Rising trends in truancy, crime and illicit drug use by young people constitute one of the most important social developments of the post-war world. One of the disturbing aspects of this development is the trend towards earlier onset of these patterns of behaviour (see Stratford and Roth, 1999 and Flood-Page *et. al.,* 2000 for evidence on and discussion in the UK context). Policy developments have been influenced by these alarming trends. In Britain, the government has adopted an ambitious target of reducing the availability and use by young people of certain types of drug by 25% by 2005 and 50% by 2008 (UKADCU, 2000). Effective anti-drugs policy may need to go beyond general targeting of this kind to much more specific action. If there is indeed a 'slippery slope' from early minor offending through soft drugs to hard drugs and serious crime, then we need to ask whether there are critical stages in this causal chain, against which policy is best directed.

All such policy initiatives are presently based on limited knowledge of the behaviour underlying these trends. It is not easy to study these issues. Illicit behaviour is inherently difficult to observe by means of conventional survey instruments. Even with suitable data, it is hard to resolve the dynamic causal structure underlying observed sequences of initiation to different types of offending and drug use because of the pervasive role of common unobservable psychological and social factors. Although there is a substantial research literature dealing with the dynamics of drug use and criminal activity at the individual level (see Flood-Page *et. al.* (2000) and Kenkel *et. al.* (2001) for), few studies concentrate explicitly on the age of initiation into crime and drug use and on the sequences in which these initiation events occur. This paper uses recent British youth survey data to examine the pathways along which early drug/crime careers evolve.

2 The 1998 Youth Lifestyles Survey

The 1998 YLS covers the 12-30 age group, who were identified through one or other of two methods. A core sample of 3643 young people was identified from households participating in the 1998 British Crime Survey (BCS) and then topped up by screening the occupants of addresses adjacent to those of the core sample to identify further subjects in the target age group. The

top-up sample was biased towards areas with high victimisation rates. Fieldwork took place during Oct 1998-Jan 1999, with Computer Assisted Self Interviewing used for the sensitive topics of drug use and criminal activity. The response rate of 69.1% yielded a usable sample of 3901 respondents. See Stratford and Roth (1999) and Flood-Page *et. al.* (2000) for further detail. The questionnaire gives information on current and past behaviour and past (at age 15) family circumstances (current circumstances for under-16s). A summary of the variables is given in Pudney (2000a).

Initially, we use a detailed breakdown of drug use into 12 illicit substances.[1] The questionnaire asks for the age at which each of these substances was first consumed. The drug with earliest onset, around age 14, is glue/solvents (also alcohol and tobacco). Over three-quarters of people reporting experience of these substances started before 16. There is then a gap of around 2bd years before the mean age of first use of cannabis and amyl nitrite. A little later, at age 17-18, comes the first use of hard drugs (heroin and crack) and other substances (amphetamines, LSD, mushrooms, tranquilisers). The most 'adult' drugs are methadone, ecstasy and finally cocaine, with a mean age of almost 20. There seems to be a natural division of drugs into five groups: (i) early onset legal substances (alcohol, tobacco); (ii) glue/solvents; (iii) early/middle onset soft drugs (amphetamines, cannabis, LSD, mushrooms, tranquilisers, amyl nitrite); (iv) early/middle onset hard drugs (heroin, crack, methadone); (v) late onset recreational drugs (ecstasy, cocaine). Crime is represented by two groups of offences. The first is a set of 18 'minor' offences (criminal damage, arson, theft, dealing in stolen goods, cheque and credit card offences, fraud and public fighting) and 9 'serious' crimes (theft of vehicles, robbery, breaking and entering and assault); see Flood-Page *et. al.* (2000, appendix B). There is a progression from truancy to minor crime to serious crime, occuring early relative to most drug use.

The sequencing of truancy/crime/drug use events is summarised in Table 1 by weighted sample frequencies of the logically possible event sequences. Two alternative definitions of crime and drug use are used: (i) all drugs (excluding alcohol and tobacco) and all crime; (ii) only hard drugs (cocaine, crack, heroin, methadone) and serious crime. There is a tendency towards

[1]The YLS also contains questions about a non-existent drug "semeron". Extremely few respondents claim experience of semeron and they have been dropped. See Pudney (2000a) for a summary of average ages of first use.

a chain of events beginning with petty crime and truancy, and only later developing into drug use. Sequences of offending beginning with drug use have a significantly smaller sample frequency than sequences beginning with truancy or crime, and this is particularly true when we consider only hard drugs and serious crime. If we were prepared to assume that this tendency has causal significance then we might conclude that a policy addressing truancy and other problems at school might be more effective than a policy attacking drug use directly. We now examine this issue by estimating conditional models of drug use and offending behaviour.

Table 1 Sequences of illicit behaviour

Sequence	**% frequency (all crimes and drugs)**	**% frequency (serious crimes and hard drugs)**
No offending or drug use	34.86	61.76
Truancy only	6.33	22.96
Crime only	11.55	3.24
Drugs only	8.83	2.57
Truancy→drugs	4.65	3.33
Truancy→crime	2.76	3.08
Crime→drugs	8.82	0.14
Crime→truancy	2.48	1.57
Drugs→truancy	1.67	0.10
Drugs→crime	5.85	0.20
Truancy→crime→drugs	5.48	1.50
Truancy→drugs→crime	5.06	0.53
Crime→truancy→drugs	6.39	0.49
Drugs→crime→truancy	1.84	0.14
Drugs→truancy→crime	2.49	0.13
Crime→drugs→truancy	3.38	0.11

Note: tied events are double-counted; alcohol and cigarettes are not included in drug use

3 Availability and demonstration effects

Economists tend to emphasise individual behaviour in isolation from the social context. The theory of rational addiction (Becker and Murphy, 1988; Grossman and Chaloupka, 1998; Kenkel, Mathios and Pacula, 2001) is an example: drug users are seen as rational individuals pursuing a planned course of action taking account of possible future consequences of current actions. It is easy to ridicule this approach when applied to behaviour that may involve severe distress and departure from normal psychological and social functioning. Nevertheless, the core idea of rational individual choice is an important one that has a place in the study of drug use. Equally, it is important to take account of social externalities (corresponding to the idea of a 'drug culture'). Manski (2000) discusses the importance of and analytical difficulties raised by these social interactions.

To illustrate the impact of external influences in an economic model of drug use, consider the following generic demand model:

$$\delta = \delta(p, u, \varphi\Lambda) \tag{1}$$

where δ is the individual's demand for the illicit drug, p is its price, u a variable distinguishing the different types of individual in the population, Λ represents the external effects influencing individuals of type u and φ is a non-negative parameter governing responsiveness to these influences. The function δ satisfies $\delta_p < 0$, $\delta_\varphi \equiv \partial\delta/\partial(\varphi\Lambda) > 0$. The vector u, which has population distribution $G(u)$, may include elements such as income, location and psychological characteristics such as risk aversion or ability.

The phenomena represented by Λ include social externalities (demonstration effects, peer pressure, etc.) and local availability through the medium of drug-using social contacts. The externality function Λ is defined implicitly by:

$$\Lambda(p, u, \varphi) = \int_{\mathsf{Z}} \delta(p, v, \varphi\Lambda(p, v, \varphi))\theta(u, v)dG(v) \tag{2}$$

where $\theta(u, v)$ is a non-negative measure of social distance or of the influence exerted by someone of type v on someone of type u. In Manski's (1993, 2000) terminology, (1) and (2) embody the notion of endogenous interactions, where each person's behaviour depends on that of the others. Contextual and correlated effects are also captured in this framework through the variable u.

A social demand equilibrium is a pair of functions $\delta(.)$, $\Lambda(.)$ satisfying (1)-(2). Aggregate demand is:

$$D(p,\varphi) = \int \delta(p,u,\varphi\Lambda(p,u,\varphi))dG(u) \tag{3}$$

Let the aggregate supply function be $S(p,t)$ with S_p, $S_t > 0$ where t represents an autonomous driver of supply. In equilibrium $D(p,\varphi) = S(p,t)$. For aggregate demand, omitting unnecessary arguments and using subscripts to denote partial derivatives:

$$D_p = \int \left[\delta_p(u) + \varphi\delta_\varphi(u)\Lambda_p(p,u,\varphi)\right] dG(u) \tag{4}$$

$$D_\varphi = \int \left[\delta_{\varphi\Lambda}(u) + \delta_\varphi(u)\left(\Lambda(p,u,\varphi) + \varphi\Lambda_\varphi(p,u,\varphi)\right)\right] dG(u) \tag{5}$$

Make the reasonable assumption that individual demands are downward-sloping after allowing for external effects. Then $\Lambda_p(p,u,\varphi) < 0$ for all v and thus $D_p < 0$. Assume that increasing the strength of the demonstration effect has a positive impact on individual demands: $\Lambda_\varphi(p,u,\varphi) > 0$ for all v and thus $D_\varphi > 0$. From (4), the effect of the social externality is to increase the price elasticity of demand. Social externalities then amplify the effect of price changes: as price falls, there is a direct increase in the individual's demand, but the consequent general increase in consumption also strengthens the demonstration effect, which stimulates individual demand still further. This process continues until a new equilibrium is reached. The effect of an expansion in supply is therefore to reduce the fall in price and increase the rise in consumption that would otherwise occur. If social interactions exist, they have the consequence of reducing the effectiveness of supply-disruption policies, so empirical evidence on their magnitude is important.

There are serious identification issues to be overcome in modelling the impact of social externalities (the reflection problem of Manski 1993). Since group norms are defined as averages of individual outcomes, it is impossible to distinguish the effect of that average from the other factors underpinning those outcomes. In dynamic models with lags in the formation of group norms, the set of possible observable outcomes is richer and the identification problem is resolved (Manski, 1993; Brock and Durlauf, 2000). This dynamic setting is inherent to the behaviour studied here. We do not pursue the identification issue formally, but rely on the dynamics of the drug use process to justify the approach used here.

It is difficult to capture empirically the influence of social externalities without having detailed information on the behaviour of subjects' social contacts. Instead, we use a macro-level proxy for the external influences acting on the individual at the time when s/he is of age t:

$$\Lambda_t(p_t, u, \varphi) \approx f(z, A_t) \tag{6}$$

where z is a set of observable variables governing the individual's social location and A_t is a macro-level index of drug use in society at large at the time that the individual is at age t. Suitable indices of prevalence have been constructed by Pudney (2001b) for a subset of the drugs considered here, using a time-series latent variable approach involving multiple indicators (drug seizures; numbers of new addicts; numbers of drug-related convictions; and BCS prevalence rates). The aggregates are plotted in Figure 1 and in some cases (notably amphetamines, LSD and cannabis) follow a path that would be difficult to capture using simple time trends. In the econometric modelling discussed later, the relevant prevalence index is used in log form.[2]

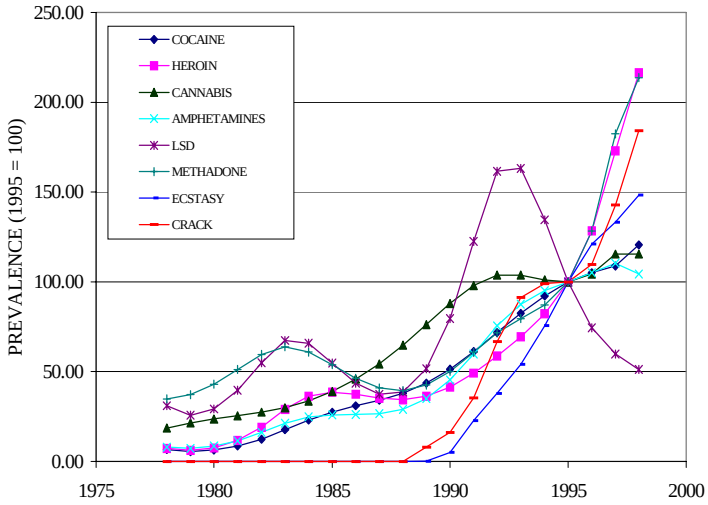

Figure 1 Indices of prevalence by drug type (source: Pudney, 2001)

[2]For crack and ecstasy, we assume a prevalence of 0.5% of the 1995 level for the period prior to 1989, during which the recorded indicators of drug use were non-existent or too low to permit a positive estimate of prevalence.

Price effects raise difficult measurement problems. The available data on street prices of illicit drugs are sparse and not very reliable. They fall far short of the quality of a conventional price index and are only available in anything like consistent form for the period since 1988. The main source is the UK National Criminal Intelligence Service, which gives rough ranges of typical street prices in a few particular locations.[3] To incorporate price effects explicitly, we would require a sequence of past prices covering the relevant past of people aged up to 30 in 1998. This would involve price series going back to 1978 but only half of that period is available. Moreover, given the inherent unreliability of price data, one can infer from available data little more than a fairly steady downward trend in the real price of the major illicit drugs over the 1988-98 period of roughly 3% per year. It would be rash to attribute much significance to the large year-to-year or between-area price variations. A further issue is supply constraints. Drugs are illicit commodities, not routinely available to everyone. It is likely that many individuals in the YLS sample will have been supply constrained for significant periods. This is particularly important in the early part of drug use careers, which are our concern. Given the incomplete and unreliable price data and the unobserved but widespread quantity constraints on demand, there is little point in attempting a standard type of demand analysis with explicit use of price variables. Instead, we rely on the constructed prevalence indices to act as proxies for consumption externalities, availability and also price movements.

4 Sequential modelling

Table 1 suggests the possibility of extensive dynamic links between types of illicit behaviour. It leaves open the possibility of causal chains going from truancy, etc. to soft drug use and on to hard drugs and serious crime. However, these links may have no causal significance and might stem from the common effect of unobserved psychological and sociological characteristics. If an individual is predisposed towards illicit behaviours by some personal characteristic, then there may be a tendency to observe involvement in truancy, crime and drug use even without any direct causal connection between them. The

[3]See Pudney (2001a) for more detail.

unobservable characteristics underlying these spurious associations might include such features as a disturbed family background, an under-developed ability to appreciate the long-term consequences of current actions, or low ability leading to under-achievement and alienation. Modelling the effect of unobservable characteristics is inherently difficult and can only generate clear results under strong assumptions. Nevertheless, it is worth attempting since the results can give an indication of the potential importance of unobservable factors and the robustness of results from simpler models.

Our analysis differs from the influential recent work of Fergusson and Horwood (2000) who used prior experience of cannabis as an explanatory factor in a proportional hazards duration model of the age of onset of use of other illicit drugs. Firstly, the risk of onset of drug use and other types of offending varies greatly with age in a non-monotonic fashion. Many of the widely-used parametric survival models, such as the Weibull used by Fergusson and Horwood may therefore be inappropriate. Secondly, age of onset is recorded only as an integer, with the consequence that there are very large numbers of 'tied' durations, causing difficulties for the alternative semi-parametric Cox regression model. Thirdly, the rapidly-changing prevalence of drug use introduces time-varying covariates linked to calendar time rather than age. Finally, there may be persistent unobservable individual-specific effects which complicate the problem of inferring causal processes from observed drug use/offending histories. The more ambitious analysis of van Ours (2001), dealing with the dynamic interrelation between alcohol, tobacco, cannabis and cocaine, uses a heterogeneous continuous-time duration model. Our approach is a little different. We analyse a wider range of behaviours, including crime as well as drug use. We also consider only the timing of onset and leave aside the issue of exit from these behavioural patterns.[4]

For any given individual, consider an observation period that covers the years from an initial age T_0 (= 11) to the current observed age T_1. There are J different types of events: the first use of each of the set of different drugs; the first episode of truancy and the first criminal offence of two types: minor and serious. Denote the ages at which these events occur by $\tau_1 ... \tau_J$.

[4]It is difficult both conceptually, and within the YLS questionnaire structure, to define a date of exit from crime or drug use. For many people, these are infrequent activities and there is no obvious date at which they can be said to have been 'given up'.

If event j is not observed within the observation period, then τ_j is censored at the arbitrary value $T_1 + 1$. We assume that the occurrences of events $1...J$ are contemporaneously independent conditional on $\{\mathbf{x}_{jt}, u_j\}$. This still permits considerable dependence through lagged effects embodied in $\mathbf{x}_{jt}$ and through correlation in the joint distribution of $u_1 .u_J$. The probability of the observed joint event $(\tau_1...\tau_J)$ is $\Pr(\tau_1...\tau_J|\mathbf{X}) = \int \mu(u_j)dG(\mathbf{u})$ where $\mu(\mathbf{u})$ is the conditional probability $\Pr(\tau_1...\tau_J|\mathbf{X}, \mathbf{u})$:

$$\mu(\mathbf{u}) = \prod_{j=1}^{J} \left[\prod_{t=1}^{\tau_j - 1} \left(1 - \Phi(\mathbf{x}_{jt}\boldsymbol{\beta}_j + u_j)\right) \Phi(\mathbf{x}_{j\tau_j}\boldsymbol{\beta}_j + u_j)^{d_j} \right] \tag{7}$$

where $\mathbf{X} = \{\mathbf{x}_{jt}\,,\ j = 1...J;\ t = 1...T\}$. We allow the random effects $u_1...u_J$ to have different variances and to be cross-correlated, by expressing the u_j as linear combinations of a set of underlying independent standardised variates:

$$\mathbf{u} = \mathbf{R}\boldsymbol{\varepsilon} \tag{8}$$

where $\mathbf{R}$ is a $J \times J$ loading matrix subject to a set of $J(J-1)/2$ restrictions normalising $\mathbf{R}$ to be a lower-triangular matrix.[5] The parameters are estimated by maximising an objective function based on a second-order expansion of the log simulated likelihood function.[6] This SML estimator is consistent and asymptotically normal with covariance matrix given by the usual inverse Hessian expression provided Q goes to infinity at least as fast as n. We use $Q = 50$ replications in our calculations, which experience with similar models suggests is adequate to make SML approximate true ML adequately (see Mealli and Pudney, 1996).

Joint modelling of this kind encounters the curse of dimensionality, since the number of parameters rises quadratically with J. We use a simplified structure with $J = 6$. Even so, an unrestricted model would involve 156 parameters. The construction of the aggregate categories is based loosely on mean ages of onset and the results of a preliminary single equation analysis

[5] We do not give a formal analysis of the identifiability of this model. However, the theoretical results of Abbring and van den Berg (2000) indicate that the identification of endogenous treatment effects is considerably less problematic in a duration setting than in the usual 2-period discrete setting. In general, in their bivariate framework, identification is achievable without the exclusion restrictions required in the conventional selection models.

[6] See Pudney (2000a) and Gouriéroux and Monfort 1996, page 45 for details, but note the minor error in equation 3.4 in the latter.

(Pudney 2000a). We distinguish glue/solvents from other soft drugs because of its early age of onset and thus its potentially important role in initiation to drug use. Ecstasy and cocaine are included in a single category because of their relatively high age of onset and their much more socially acceptable image than heroin and crack. The six categories finally specified are: (i) *glue/solvent abuse*; (ii) *soft drug use* (cannabis, amphetamines, LSD, magic mushrooms, amyl nitrite); (iii) *'social' drug use* (ecstasy or cocaine); (iv) *hard drug use* (heroin, crack, methadone); (v) *minor offending* (truancy, criminal damage, arson, theft, dealing in stolen goods, cheque and credit card offences, fraud and public fighting); (vi) *serious crime* (theft of vehicles, robbery, burglary and assault). The SML estimator for this model is computationally demanding. Our strategy is to begin with the 1-factor model in which $\boldsymbol{\varepsilon}$ contains a single random factor and the matrix $\mathbf{R}$ is a column vector. Then a sequence of generalised models is estimated, with the number of random factors in $\boldsymbol{\varepsilon}$ increased by 1 at each step. This process is terminated when the addition of an extra factor leads to an insignificant improvement according to a simulated likelihood ratio criterion. By this criterion, the 1-factor results dominated the 2-factor and are the ones discussed below.[7] Full parameter estimates are given in Pudney (2000a).

The YLS is not a panel survey, so characteristics summarising family background are only observable at one point in time. These variables record whether parents were absent from the family and their employment status at the time of the respondent's fifteenth year of age or survey date, whichever is earlier. Other variables describing the neighbourhood (inner city and/or socially deprived); any family history of trouble with the police; and any religious affiliation, are observable only at the time of interview. We include the relevant (log) prevalence variable as a covariate in the equations for 'soft' drugs, ecstasy/cocaine and 'hard' drugs. For the remaining categories, a quadratic time trend approximates the effect of changing conditions over time. In every case where it is available, the use of the prevalence variable resulted in a better fit that the time trend. Full parameter estimates for this model are given in Pudney (2000a). The estimated parameters for

[7] A comparison of the likelihood values computed at a representative point confirmed that the simulation approach delivers numerical accuracy comparable with Gauss-Legendre quadrature (see Butler and Moffitt, 1982). Treating 40-point quadrature as fully accurate, the simulated likelihood with $\mathtt{Q} = 50$ gave a similar degree of accuracy to 20-point quadrature. Standard software often uses quadrature based on as few as 12 points.

the effects of past behaviour turn out to be generally smaller than in the non-heterogeneous variant of the model. There is no significant impact of early soft drug use on the hazard for hard drugs (crack, heroin, methadone), although there is still a significant impact of soft drugs on the hazard for truancy and minor crime. In contrast, the estimated impact of social and family background and of general drug prevalence is stronger in the random effects model.

5 The impact of personal characteristics

We use stochastic simulation to assess the impact of personal characteristics on behaviour, examining four variations on a hypothetical baseline individual type: a white male with a favourable family/social background (both parents present and in work, no family history of trouble with police, resident in non-deprived non-inner-city area) and living through a stable period with low prevalence of drug use (10% soft drugs; 0.5% ecstasy/cocaine; 0.1% hard drugs). The other four hypothetical individuals are deviations from this base: (i) a disadvantaged background (absent non-working father, working mother, family history of trouble with police, resident in deprived inner-city area); (ii) female; (iii) Asian; (iv) strong drug culture (50% of the population having ever used soft drugs, 8% ecstasy/cocaine and 2% hard drugs). Three sets of simulations are presented. Table 2 is based on estimates of the non-heterogeneous model (with $\mathbf{R}$ constrained to zero), while Tables 3 and 4 are based on the 1-factor random effects model. The simulation in Table 3 holds the random effects $\mathbf{u}$ fixed at zero, while those in Table 4 allow $\mathbf{u}$ to vary randomly across replications, representing a cross-section of individuals of individuals with the same observable characteristics but differing predispositions. Simulations are done by generating 50,000 6×18 pseudo-random matrices (6 drug/offending categories and 18 years from age 12 to 29). From each of these we generate a personal history for each of the hypothetical individual types. In Table 4, the random effects $\mathbf{u}$ are sampled from the distribution $G(\mathbf{u}) = N(\mathbf{0}, \mathbf{R}\mathbf{R}')$. For the baseline, the tables give the proportion ($\hat{\boldsymbol{p}}_j$) of the replications yielding experience of drug or offence j and the average age of onset ($\overline{\tau}_j$) in those cases. For the other cases, figures quoted are the differences from the baseline.

A disadvantaged social/family background is clearly the dominant influ-

ence on drug use and offending, with drug culture (as measured by prevalence) also extremely important. With the exception of a small but statistically significant rise in the hazard rate for serious crime for blacks, the influence of gender and ethnicity is to reduce the incidence of drug use and offending in comparison with the baseline white male group. The ethnicity results give no support for the common racial stereotype of the young black criminal-drug user.

Table 2 Predicted % prevalence ($\hat{p}_j$) and mean age of onset ($\overline{\tau}_j$) for baseline individual and differences ($\Delta\hat{p}_j$, $\Delta\overline{\tau}_j$) relative to the baseline (model without random effects)

		Solv.	Soft	C & E	Hard	Minor	Serious
Baseline	$\hat{p}_j$	4.6	46.2	13.5	1.6	68.6	9.2
white male	$\overline{\tau}_j$	14.4	17.0	19.8	19.0	14.6	15.2
Disadvant-	$\Delta\hat{p}_j$	+26.1	+48.4	+28.6	+25.3	+31.1	+55.3
aged	$\Delta\overline{\tau}_j$	+0.4	-1.6	-0.5	+0.1	-2.1	+0.1
Female	$\Delta\hat{p}_j$	-1.7	-14.0	-8.1	-0.9	-16.7	-7.5
	$\Delta\overline{\tau}_j$	-0.2	+0.1	-0.1	-1.0	+0.2	-0.5
Asian	$\Delta\hat{p}_j$	-2.2	-34.0	-10.7	+0.5	-20.7	-4.1
	$\Delta\overline{\tau}_j$	-0.3	+0.3	-0.4	-2.0	+0.0	-0.7
Black	$\Delta\hat{p}_j$	-3.7	-19.2	-8.5	-1.2	-2.2	4.8
	$\Delta\overline{\tau}_j$	-0.0	+0.2	-0.0	-1.0	-0.1	-0.3
High	$\Delta\hat{p}_j$	+10.6	+49.0	+43.5	+34.9	+19.2	+22.8
prevalence	$\Delta\overline{\tau}_j$	+0.5	-1.8	-0.8	-0.0	-0.1	+0.9

Table 2 appears to support an indirect drugs policy directed initially at reducing social exclusion and disadvantage. However, causation is difficult to establish in this non-experimental setting and it is important to explore the robustness of these results to specification changes which allow for non-causal association. The simulations of Tables 3 and 4 use the random effects model and thus allow for unobservable factors representing a predisposition of some individuals towards patterns of drug use and offending. The prevalence of drug use and offending among the simulated individuals is generally much higher in Table 4 (allowing for random variation in $\mathbf{u}$) than in Table 3 (with

u fixed at **0**). Thus drug use appears partly the result of intrinsic person-specific factors deviating from the statistical norm. There is a fair degree of robustness in the unconditional predictions of the heterogeneous and non-heterogeneous models (compare Table 2 with Table 4), contrasting with the lack of robustness in the conditional causal predictions (compare Table 2 with Table 3). For the heterogeneous model, the simulated 'causal' influence of these variables can be found by making conditional simulations, holding **u** fixed at **0** (Table 3). The predicted impact of social disadvantage and general drug prevalence on the hazard rates for soft and 'social' drugs (cannabis, amphetamines, etc. and cocaine/ecstasy) become considerably larger, while their predicted impact for 'hard' drugs (heroin, crack, methadone) is reduced. This implies that individuals of statistically 'abnormal' type (in terms of their value of **u**) have a hard drug hazard that is more sensitive to adverse social/family conditions than are more 'normal' types. The reverse applies for softer drugs.

Table 3 Predicted % prevalence ($\hat{\mathbf{p}}_j$) and mean age of onset ($\overline{\tau}_j$) for baseline individual and differences ($\Delta\hat{\mathbf{p}}_j$, $\Delta\overline{\tau}_j$) relative to the baseline (random effects model; **u** set to **0**)

		Solv.	Soft	C & E	Hard	Minor	Serious
Baseline	$\hat{\mathbf{p}}_j$	1.2	36.6	1.7	0.0	67.1	5.5
white male	$\overline{\tau}_j$	14.7	18.2	20.6	16.7	14.8	15.2
Disadvant-	$\Delta\hat{\mathbf{p}}_j$	+25.5	+63.3	+62.4	+13.9	+32.9	+60.9
aged	$\Delta\overline{\tau}_j$	-0.3	-3.6	-1.4	+2.5	-2.7	-0.4
Female	$\Delta\hat{\mathbf{p}}_j$	-0.6	-16.9	-1.5	-0.0	-19.1	-4.8
	$\Delta\overline{\tau}_j$	-0.2	+0.1	-0.2	+0.3	+0.2	-0.6
Asian	$\Delta\hat{\mathbf{p}}_j$	-0.7	-33.4	-1.7	+0.0	-23.4	-3.3
	$\Delta\overline{\tau}_j$	-0.1	+0.1	-0.8	+0.6	+0.2	-0.5
Black	$\Delta\hat{\mathbf{p}}_j$	-1.0	-21.7	-1.6	-0.0	-1.7	+2.8
	$\Delta\overline{\tau}_j$	+0.1	+0.2	-0.7	-16.7	-0.0	+0.1
High	$\Delta\hat{\mathbf{p}}_j$	+1.4	+60.6	+17.0	+3.8	+13.5	+11.2
prevalence	$\Delta\overline{\tau}_j$	-0.1	-2.2	-0.3	+2.1	-0.2	+0.2

Table 4 Predicted % prevalence ($\hat{P}_j$) and mean age of onset ($\overline{\tau}_j$) for baseline individual and differences ($\Delta\hat{P}_j$, $\Delta\overline{\tau}_j$) relative to the baseline (random effects model; $\mathbf{u}$ generated as $N(\mathbf{0}, \mathbf{I})$)

		Solv.	Soft	C & E	Hard	Minor	Serious
Baseline	$\hat{P}_j$	7.0	43.7	13.1	2.6	64.7	9.9
white male	$\overline{\tau}_j$	14.3	16.9	19.3	18.5	14.5	15.2
Disadvant-	$\Delta\hat{P}_j$	+27.5	+48.0	+43.6	+28.3	+34.5	+52.7
aged	$\Delta\overline{\tau}_j$	-0.6	-2.3	-1.7	-1.7	-2.2	-0.9
Female	$\Delta\hat{P}_j$	-2.5	-11.1	-6.8	-1.4	-15.1	-7.9
	$\Delta\overline{\tau}_j$	+0.1	+0.3	+0.4	+0.1	+0.3	-0.0
Asian	$\Delta\hat{P}_j$	-2.4	-30.0	-10.1	+0.4	-18.9	-4.7
	$\Delta\overline{\tau}_j$	+0.2	+0.9	+0.5	-0.5	+0.3	-0.1
Black	$\Delta\hat{P}_j$	-4.5	-15.5	-9.1	-2.0	-1.2	+3.9
	$\Delta\overline{\tau}_j$	+0.2	+0.5	+0.4	+0.1	+0.0	+0.0
High	$\Delta\hat{P}_j$	+2.7	+38.4	+17.5	+15.4	+10.0	+12.5
prevalence	$\Delta\overline{\tau}_j$	-0.1	-1.5	-0.7	-1.1	-0.2	-0.2

6 The gateway effect

The gateway effect is the increase in hazard rate for onset of hard drug use induced by prior use of soft drugs after controlling for observed and unobserved characteristics. Figures 2-6 illustrate the estimated gateway effect, showing the abstention or survivor probability $\Pr(\tau_j > t|\mathbf{x}_{j1}...\mathbf{x}_{jt}, u_j) = \prod_{s=1}^{t} 1 - \Phi(\mathbf{x}_{js}\boldsymbol{\beta}_j + u_j)$ plotted against age t. Each panel shows four cases: (i) the baseline white male in favourable circumstances and no previous experience of drugs or offending; (ii) similar, except for early (age 12) experience of the relevant drug or offence; (iii) a white male with absent father and working mother, living in a deprived inner-city area at a time of high drug prevalence but no personal history of drugs or offending; and (iv) similar but with early experience of the relevant drug or offence. There are thirty possible plots corresponding to the effect of each of the six drugs/offences on the other five. We show only the five for which there is a significant positive gateway coefficient in the random effects model. The preferred random effects specification is used, with u_j set to zero.

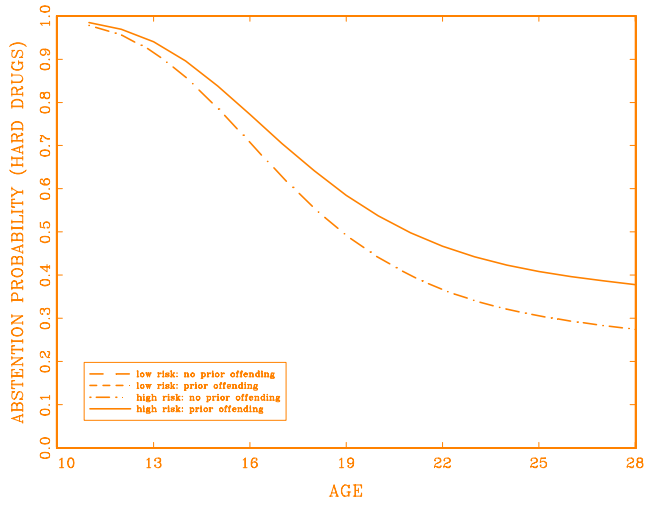

Figure 2 The effect of prior truancy/minor crime on the abstention probability for hard drugs

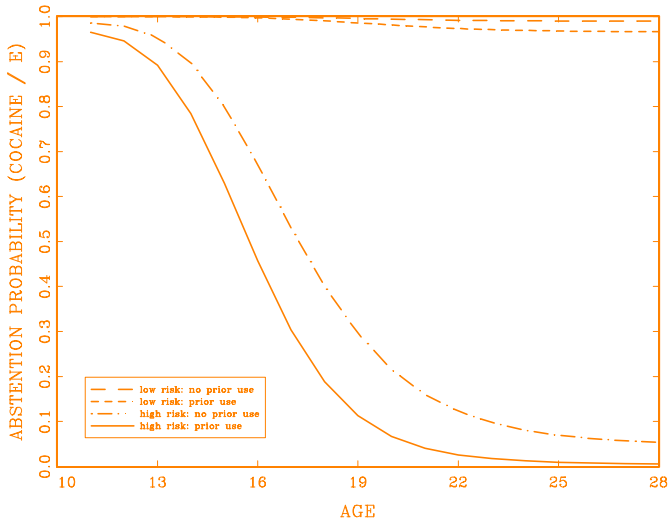

Figure 3 The effect of prior soft drug use on the abstention probability for cocaine/E

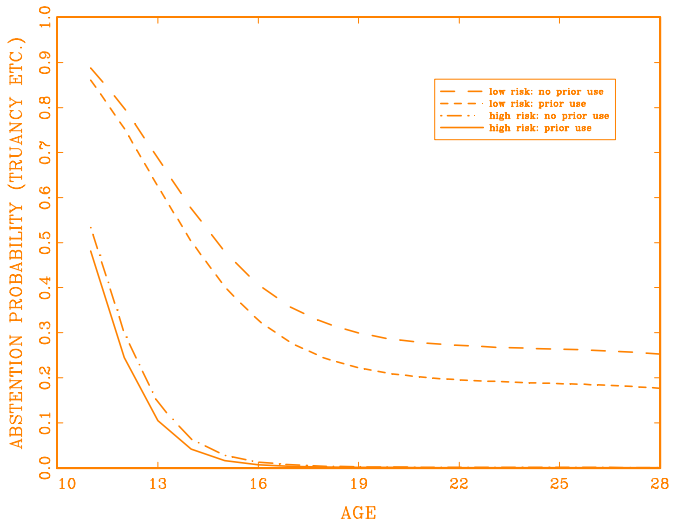

Figure 4 The effect of prior soft drug use on the abstention probability for truancy/minor crime

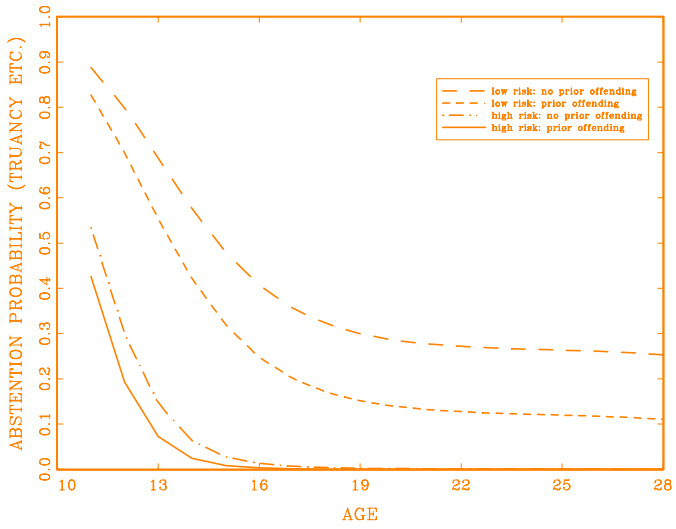

Figure 5 The effect of prior serious crime on the abstention probability for truancy/minor crime

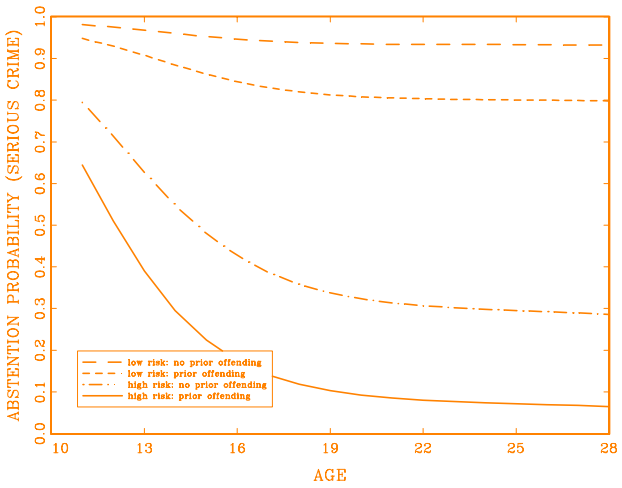

Figure 6 The effect of prior truancy/minor crime on the abstention probability for serious crime

Note that, in Figure 2 the abstention probability is essentially unity for hard drugs in the low-risk case, so only one pair of curves appears. The most substantial gateway effects are for the two categories of crime into hard drugs and other crime (Figures 2, 5, 6). There is evidence of a smaller gateway from soft drug use into ecstasy/cocaine and minor crime(Figures 3, 4), but these effects are small compared to those of age and social background. [8]

Figures 2-6 ignore possible indirect effects of the use of one drug on subsequent use of another via intermediate drug types. Table 5 gives a fuller picture using stochastic simulation. We first simulate a complete offending history for the age range 11-30 for each individual in the YLS sample This is then repeated (re-using the same set of pseudo-random variates), but with the hazard rate for one of the drug categories constrained to be zero. The result is then compared with the baseline. The comparison is done for each drug/offending category in turn. The random-effects model is used and, to make the results representative of the YLS target population, we use a fresh draw of the random effect $\mathbf{u}$ for each individual (held constant across all simulations for that individual). The comparisons give an assessment of the

[8]Pudney (2000a) compares these with analogous plots for the non-hereogeneous model. This reveals that selection bias greatly exaggerates the gateway effect.

causal impact, or gateway effect, of each type of drug use/offending on others after controlling for the influence of unobservables. The gateway effects turn out to be moderate.[9] Although there are statistically significant positive effects in 8 of the 30 possible cases, these impacts are small and there is little convincing evidence of quantitatively important causal pathways leading from one type of drug use to others. According to these results, if we were somehow to find a policy that would totally remove soft drugs from the scene, the largest statistically significant impact would be a reduction of around one third in the prevalence of cocaine / ecstasy use. Hard drug use would rise by an (insignificant) 8%.

Table 5 Impact of early experience of solvent abuse, soft drugs, truancy or crime on subsequent behaviour

Effect on use of ...		Solv.	Soft	E/coc	Hard	Tru.	S.crime
Baseline case	$\hat{p}_j$	8.2	40.2	11.7	2.5	55.9	9.1
	$\overline{\tau}_j$	14.1	16.7	19.2	17.9	14.0	14.9
Effect of removal of risk arising from ...							
...Glue/solvents	$\Delta\hat{p}_j$	-	-0.2	+1.1	+0.6	+0.0	+0.1
	$\Delta\overline{\tau}_j$	-	+0.0	-0.2	+0.2	+0.0	+0.0
...Soft drugs	$\Delta\hat{p}_j$	+0.7	-	-3.8	-0.2	-0.6	+0.2
	$\Delta\overline{\tau}_j$	+0.2	-	-0.5	-0.4	-0.0	+0.1
...Ecstasy/cocaine	$\Delta\hat{p}_j$	+0.2	+0.0	-	-0.3	+0.1	+0.0
	$\Delta\overline{\tau}_j$	+0.0	-0.0	-	-0.6	+0.0	+0.0
...Hard drugs	$\Delta\hat{p}_j$	+0.0	+0.1	+0.2	-	+0.0	+0.1
	$\Delta\overline{\tau}_j$	+0.0	-0.0	-0.0	-	-0.0	+0.1
...Truancy, minor	$\Delta\hat{p}_j$	-2.0	-1.3	+0.9	+0.8	-	-3.3
crime	$\Delta\overline{\tau}_j$	-0.2	+0.0	-0.0	+0.1	-	-1.0
...Serious crime	$\Delta\hat{p}_j$	-0.2	+0.2	+0.3	+0.1	-0.2	-
	$\Delta\overline{\tau}_j$	+0.0	+0.0	-0.1	+0.2	-0.0	-

[9]For the non-heterogeneous model, estimated without random effects, analogous simulations indicate misleading large gateway effects for truancy and minor crime, with particularly large reductions in the prevalence of all types of drug use (see Pudney 2001a).

7 Conclusions

We have applied discrete statistical duration methods to data from the 1998 Youth Lifestyles Survey to investigate the age of onset of various types of crime and illicit drug use. The YLS data suggest at first sight that there are widespread and strong links between early experience of minor offending/drug use and later involvement in more serious crime and drug abuse. However, this is not a robust finding. After allowing for unobservable individual-specific random effects, the estimates of these dynamic impacts are reduced considerably and remain small even where statistically significant. This is in line with the results of analogous work by van Ours (2001) using data from Amsterdam. The estimated impacts of social, family and cultural factors are much more important. The background 'drug culture' of society, as proxied by aggregate drug prevalence trends, is also a dominant influence.

The policy implications of these findings are important. There is little support for a policy directed at reducing exposure to soft drugs. Our best-fitting statistical model suggests that the impact of even the most effective anti-soft drug policy imaginable would not make a noticeable difference to hard drug use. Our results suggest instead that an effective policy directed at reducing the extent of social deprivation may have a better chance of success.

References

[1] Abbring, J. H. and van den Berg, G. J. (2000). The non-parametric identification of treatment effects in duration models. Working paper, Free University Amsterdam.

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

[3] Brock, W. and Durlauf, S. (2000). Interactions-based models, in *Handbook of Econometrics, vol. 5*, J. Heckman and E. Leamer (eds.). Amsterdam: North-Holland.

[4] Butler, J. S. and Moffitt, R. (1982), "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model", *Econometrica*, 50, 761-764.

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

[6] Flood-Page, C., Campbell, S., Harrington, V. and Miller, J. (2000). *Youth Crime: Findings from the 1998/99 Youth Lifestyles Survey*. London: Home Office Research Study no. 209.

[7] Gourieroux, C. and Monfort, A. (1996) *Simulation-Based Econometric Methods* (Oxford University Press, Oxford).

[8] Grossman, M. and Chaloupka, F. J. (1998). The demand for cocaine by young adults: a rational addiction approach, *Journal of Health Economics* **17**, 427-474.

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

[10] Kenkel, D., Mathios, A. D. and Pacula, R. L. (2001). Economics of youth drug use, addiction and gateway effects, *Addiction* **96**, 151-164.

[11] Manski, C. F. (1993). Identification of social effects: the reflection problem, *Review of Economic Studies* **60**, 531-542.

[12] Manski, C. F. (2000). Economic analysis of social interactions, *Journal of Economic Perspectives* **14**, 115-136.

[13] Maughan, B., Pickles, A., Rowe, R., Costello, E. and Angold, A. (2000). Developmental trajectories of aggressive and non-aggressive conduct problems, *Journal of Quantitative Criminology* **16**, 199-221.

[14] Mealli, F. and Pudney, S. E. (1996), "Occupational Pensions and Job Mobility in Britain: Estimation of a Random-Effects Competing Risks Model", *Journal of Applied Econometrics*, 11, 293-320.

[15] Pacula, R. (1997). The modelling of the gateway effect, *Health Economics* **6**, 521-524.

[16] Pudney, S. E. (2001a).The road to ruin. Sequences of drug use and offending by young people in Britain. University of Leicester: discussion Papers in Public Sector Economics no. 01/3.

[17] Pudney, S. E. (2001b).The growth of illicit drugs markets in the UK 1978-99. University of Leicester: discussion Papers in Public Sector Economics no. 01/7.

[18] Stenbacka, M., Allebeck, P. and Romelsjo, A. (1993). Initiation into drug use: the pathway from being offered drugs to trying cannabis and progression to intravenous drug abuse, *Scandinavian Journal of Social Medicine* **21**, 31-39.

[19] UKADCU (2000). United Kingdom's Anti-Drugs Co-ordinator's Second National Plan 2000/2001. London: Cabinet Office.

[20] van Ours, J. C. (2001). On the dynamics in the consumption of alcohol, tobacco, cannabis and cocaine. Working paper, Tilburg University.

[21] Stratford, N. and Roth, W. (1999). *The 1998 Youth Lifestyles Survey Technical Report*, London: National Centre for Social Research.

[22] Yamaguchi, K. and Kandel, D. B. (1984a). Patterns of drug use from adolescence to young adulthood II. Sequences of progression, *American Journal of Public Health* **74**, 668-672.

[23] Yamaguchi, K. and Kandel, D. B. (1984b). Patterns of drug use from adolescence to young adulthood III. Predictors of progression, *American Journal of Public Health* **74**, 673-681.