Macroeconomic Aspects of Substance Abuse: Diffusion, Productivity and Optimal Control

Amnon Levy

and

Frank Neri

WP 04-22

December 2004

Macroeconomic Aspects of Substance Abuse: Diffusion, Productivity and Optimal Control

Amnon Levy[*]
Economics Discipline
School of Economics and Information Systems
Faculty of Commerce
University of Wollongong
Phone:+61-2-42213658
Fax:+61-2-42213725
E-mail: Amnon_Levy@uow.edu.au

Frank Neri
Economics Discipline
School of Economics and Information Systems
Faculty of Commerce
University of Wollongong
Phone:+61-2-42214671
Fax:+61-2-42213725
E-mail: Frank_Neri@uow.edu.au

Abstract

This paper deals dynamically with macroeconomic aspects of widespread substance abuse with a reference to illicit drugs as an example. Substance-abuse impedes the productivity of the labour force and subsequently economic growth. The labour force is divided into non-using and therefore fully productive workers, a number of whom are employed by the government in drug-control activities, and drug users who are only partially productive. An efficient management of the nation's portfolio of workers is taken to be the trajectory of drug-control that maximises the present value of the stream of disposable national incomes.

---

[*] Corresponding author.

1. Introduction

The trade and consumption of abusive substances, illicit drugs in particular, is a serious social problem confronting governments in many countries today. In some economically poor countries, the use of drugs has been an integral part of the local culture and has considerably affected the allocation of time, supply of labour, income, consumption and investment at the household and aggregate levels. Perhaps in no other country is this most starkly the case than in Colombia where the drug industry has acted as a catalyst in de-legitimating the regime, has diminished trust and increased transaction costs, violence, impunity, security costs, and has promoted highly speculative investments and capital flight (cf. Thoumi, 2002). Other notable examples are Yemen where addiction, though to a relatively mild drug (Kat), is the norm, and Cambodia and Afghanistan where the economic, social and political affairs in parts of these countries are dominated by the production, marketing and consumption of opium. The trade and consumption of illicit drugs has also had deleterious effects in Mexico (cf. Chabet 2002) and Nigeria (cf. Klein 1999). Illicit drugs have also adversely affected rich countries. For example, Harwood, Fountain and Livermore (1998) estimate the cost of drug abuse and dependence in the USA in 1992 at 98 billion U.S. dollars. Using a cost-of-illness approach, Xie, Rehrn, Single, Robson and Paul (1998) estimate the economic costs of illicit drug use in Ontario, 1992, to be 489 million Canadian dollars. The estimate of Collins and Lapsley (1991) for Australia in 1988 is around 1.6 billion Australian dollars.

While almost all drugs were once freely available in many countries, recent decades have seen the introduction of progressively tougher domestic and international restrictions on the traffic and consumption of mind altering drugs in particular. Kennally (2001) argues that the prohibition restricts entry, reduces

consumer information and thus increases the market power of existing traders who use violence to enforce contracts and produce products of unknown quality. Hence, increasingly stronger enforcement efforts are likely to put upward pressure on the *full* price of illicit drugs to users. This positive relationship between user full price and enforcement has been incorporated by Behrens, Caulkins, Tragler and Feichtinger (1997) and Tragler, Caulkins and Feichtinger (2001) in optimal control analyses of the cost-minimising mix of spending on enforcement and treatment.

This paper, though not incorporating the price-raising effect of drug-control explicitly, complements the analyses of Behrens *et al*. (1997) and Tragler *et al*. (2001) by focusing concerns on aggregate production and growth and the debilitating effects of illicit drugs. More specifically we focus on the relationships between disposable national income, the decomposition of the labour force between users and nonusers, and government control efforts. We construct a macro-dynamic optimisation model to describe rational public investment of effort in controlling the spread of illicit drugs and their adverse effect on the labour force and economic growth. Drug control effort is determined so as to maximise the sum of the discounted disposable national incomes in the presence of fully productive non-using workers and less productive drug users. Drug use is modelled as a diffusion process with user numbers increasing in the existing user population but decreasing in the costly government control effort. The conceptual framework developed in our paper can also be used for analysing the growth-efficient control on other potentially harmful substances and activities such as alcohol, tobacco, junk-food products and HIV infectious behaviour. Therefore, while we refer explicitly to illicit drugs, our macro-dynamic approach is sufficiently general for dealing conceptually with the abuse of any potentially debilitating substance or activity.

The paper continues as follows. Section 2 introduces the building blocks of our macro-dynamic drug-control model where the nation's labour force is divided into non-using and therefore fully productive workers and less productive drug users. Section 3 presents and interprets the rule for drug-control effort that maximises the present value of the stream of the net national incomes stemming from the country's using and non-using workers. Two interior equilibriums are identified in section 4 for the case where the intrinsic drug-use diffusion rate is smaller than the discount rate: one with a low number of users and supported by a high-control effort and the other with a low-control effort and a high number of users. The properties of these steady states are analysed and policy conclusions are drawn. The effects of the model parameters on the number of users, the control effort and the level of disposable national income in each of the two interior steady states are analysed in section 5. Section 6 outlines two extensions of the model. The first extension distinguishes between two types of drug-control activities: prevention, which we define as law enforcement and treatment activities that deter initiation, and rehabilitation where long-term dependent users are restored to a drug-free existence. We argue that convexity of the prevention and rehabilitation cost functions ensures the optimality of mixed control activities. We show how the growth-efficient mix of prevention and rehabilitation depends on their relative effectiveness and cost as well as on the difference between the discount rate and the marginal diffusion of drug-use and its determinants. The second extension explores the effect on the growth-efficient control effort of societal conflict between users and nonusers should the former group becomes a relatively large proportion of the society. A brief summary of the paper and a detailed account of its main conclusions are given in section 7.

2. Building Blocks of the Macro-Dynamic Drug-Control Model

Our macro-dynamic model focuses on the relationship between the use of illicit drugs, labour productivity, aggregate income and the government drug-control effort. For tractability, we ignore capital, capital accumulation and technological changes and assume that labour is the sole factor of production. More specifically, our model is based on the following assumptions.

Assumption 1 (labour-force size and composition): The size of the working-age population is time invariant and equal to L,[1] of which $0 \leq N(t) \leq (1 - \boldsymbol{b})L$ people use illicit drugs, hereafter *users*, where $0 \leq \boldsymbol{b} < 1$ denotes the share of the population absolutely unsusceptible to illicit drugs.

Assumption 2 (drugs and employment): L_g members of the labour force are drug-control officers, hereafter *controllers*, employed by the government and providing a composite service of prevention and rehabilitation.[2] The remaining $L - L_g$ members are employed in the private sector. C*ontrollers* are identical and cannot be *users*.

Assumption 3 (drugs and productivity): Illicit drug use reduces productivity. If the instantaneous output of each of the privately employed $L - N(t) - L_g$ *nonusers* is y then the instantaneous output of a *user* is

$$y_n = \boldsymbol{e}y, \ 0 \leq \boldsymbol{e} < 1 \tag{1}$$

---

[1] Similarly a time-invariant population size is assumed, to simplify matters, in Skiba's (1978) optimal growth analysis. In this case, the greater the aggregate output at any given moment, the greater the benefits gained by society.

[2] The extension of the model in section 6 divides L_g into a group of controllers providing rehabilitation and a group of controllers engaged in prevention.

where, y is a positive time-invariant scalar, and e is the relative productivity of a user with $e = 0$ indicating total incapacitation and $0 < e < 1$ partial incapacitation.[3]

Assumption 4 (drug-proliferation): The net conversion of *nonusers* and *controllers* to *users* is given by the difference between a concave diffusion function (F) and a linear[4] drug-control function (dL_g)

$$\dot{N}(t) = F(N(t);(1-b)L) - dL_g(t) \tag{2}$$

where d is a positive scalar denoting the instantaneous marginal and average productivity of each *controller* in terms of the number of people prevented and rehabilitated from using illicit drugs.

Our specification of the diffusion function is based on the premise that the use of narcotics is socially contagious—as the number of *users* increases, drug using

---

[3] In an inter-temporal analysis of an *individual's* production and substance abuse it is reasonable to assume that the debilitating effect of a *constant* amount of drug consumption is likely to be decreasing: $\dot{e} > 0$ as a *user* becomes gradually accustomed to the adverse effects of drugs and develops some level of tolerance to those effects. However, in an inter-temporal analysis of *national* production and substance abuse, such as the present one, $1-e$ denotes the average loss of production capacity within the group of *users*. Due to entry and exit, the composition of this group changes over time. Entry of new *users* is driven by curiosity, temptation, conformity to peers' expectations, failure and loss. Exit is due to attrition—death of heavy and old *users*—and also to rehabilitation and maturation of others. Since the membership in the group of *users* changes continually it is difficult to justify $\dot{e} > 0$ or $\dot{e} < 0$. Moreover, due to tolerance the amount and frequency of drug-consumption among veteran *users* are likely to be greater than among newly initiated *users*. The increased drug-consumption by dependent *users* offsets the positive effect of their higher tolerance to the adverse effects of drugs on their productivity. Therefore, a conservative assumption of time in-variant average loss of personal productivity within the group of *users* is made.

[4] An alternative concave specification -- $R(L_g), R' > 0, R'' < 0$ -- reflecting diminishing marginal control requires $-l(t)R''(Lg(t)) < 2c$ for an interior solution to the maximisation problem described in section 3 to exist, where c is the private and social cost coefficient indicated in assumption 5 and l is the co-state variable of the maximisation-problem's Hamiltonian. It is also possible that the marginal control effort depends on the number of *users* ($R'(L_g;N)$). However, the effect of N on the marginal control effort is not clear *a-priori*. On the one hand, the greater the number of *users* the easier the "catch". On the other hand, a larger number of *users* might be associated with a greater resistance to governmental control effort.

becomes more socially acceptable.[5] Due to innate and acquired differences, people have varying degrees of resistance to illicit drugs. Drug use spreads gradually, but in diminishing increments, from highly susceptible people to less susceptible people. Its spread is also moderated by maturation and mortality. Formally, we assume that $F > 0$ and $F'' < 0$ for all $0 < N < (1 - b)L$. Diffusion is positive as long as the upper bound is not reached, but the marginal diffusion diminishes as the degree of resistance within the remaining group of *nonusers* rises and as the natural attrition of *users* increases. In particular, we assume $F' > 0$, reflecting a positive marginal diffusion up to a critical level $N^* < (1 - b)L$. Thereafter $F' < 0$ reflecting a negative marginal diffusion, which is due to a dominant natural attrition effect, and leading to $F = 0$ when $N = (1 - b)L$. Furthermore, in order to retain a non-negative state variable we assume that $F \le 0$ for all $N \le 0$. Recalling that $F > 0$ for all $0 < N < (1 - b)L$, then $F(0;(1-b)L) = 0$.[6] Recalling also that $dL_g \ge 0$, we merely require that the terminal value of N is non-negative ($N(\infty) \ge 0$) to ensure the non-negativity of $N(t)$ for every t in the planning horizon.[7] This requirement of terminal non-negativity relieves us from having to incorporate the state-variable constraint, $0 \le N(t) \le (1 - b)L$, into the Hamiltonian associated with the optimal-control problem presented in the following section by equations (5) and (2) and $N(\infty) \ge 0$.[8]

---

[5] There exists a considerable literature that models illicit drug use as a socially contagious activity. Diffusion models have had considerable success in describing the initial introduction and subsequent spread of new substances such as illicit drugs. See Ferrence, R. (2001) for a review.

[6] Note that the logistic diffusion function used in section 4, $F(t) = aN(t)[1 - N(t)/(1 - b)L]$ with $0 < a < 1$ denoting the intrinsic diffusion rate $N^* = 0.5(1 - b)L$, satisfies these assumptions.

[7] The underlying rationale is that once N reaches zero, it cannot increase thereafter. If N were to become negative, it could not increase later to satisfy the terminal non-negativity requirement. Thus, our terminal non-negativity restriction assures non-negativity of N throughout.

[8] See Kamien and Schwartz, 1991, p. 230, for a general discussion of this issue.

Assumption 5 (control costs): The instantaneous cost of illicit drug control is an increasing and convex function of $L_g(t)$ comprising a linear part of forgone private output (y for each *controller*) and a quadratic part that comprises the private and social costs stemming from governmental control. Consistent with Tragler *et al.* (2001) and Kennally (2001), we assume these private and social costs increase in the government's control effort as drug traders resort to more serious forms of criminal activity and violence, and drug users face greater uncertainties about supply and drug quality. In formal terms, the instantaneous costs of control are depicted as follows:

$$C(t) = yL_g(t) + cL_g(t)^2 \tag{3}$$

where c is the positive coefficient of the marginal private and social costs stemming from the crime, violence and drug-quality uncertainty accompanying the government control effort.

Assumption 6 (balanced budget and tax neutrality): At every instance the government fully finances the control effort by collecting a lump-sum tax.
Recalling that the incomes of all *nonusers* are identical and the incomes of all *users* are also identical but proportionally lower, a progressive tax scheme would subsidise *users* at the expense of *nonusers* and would affect the supply of labour. In contrast, a lump-sum tax treats *users* and *nonusers* equally and does not affect the supply of labour.

3. Growth-Efficient Drug-Control Effort

Assumptions 1-3, 5 and 6 imply that the instantaneous disposable national income (DNI); that is, gross national income net of government spending on, and private and social costs of, illicit drug-control; is given by

$$DNI(t) = [L - (1 - e)N(t) - L_g(t)]y - cL_g(t)^2. \qquad (4)$$

Recalling that the size of the working-age population (L) is assumed to be time invariant, there is no need to divide the DNIs accruing at different instances by population size for inter-temporarily assessing national economic benefits. Maximising the sum of the discounted instantaneous DNIs is equivalent, in this case, to maximising the sum of the discounted per capita DNIs. A growth-efficient drug-control effort is the trajectory of the number of *controllers* ($L_g{}^o$) that maximises the sum of the discounted instantaneous DNIs generated over an infinite planning horizon subject to the conversion equation of *nonusers* and *controllers* to *users*. That is,

$$L_g{}^o = \arg\max \int_0^\infty e^{-rt} \{[L - (1 - e)N(t) - L_g(t)]y - cL_g(t)^2\}dt \qquad (5)$$

subject to the motion equation (2) and $N(\infty) \geq 0$, and where r is the planner's positive fixed rate of time preference. The Hamiltonian associated with this decision problem is concave in the control variable (L_g). The necessary conditions for maximum and the no-arbitrage rule are derived in Appendix A. Since the co-state variable $l(t)$ multiplying $[F(N(t); (1 - b)L) - dL_g(t)]$ is non-positive and the diffusion function $F(N(t); (1 - b)L)$ is assumed to be concave in N, the Hamiltonian is non–concave in N and the Mangasarian's theorem on the sufficiency of the Pontryagin's maximum-principle conditions is not valid in this case. As speculated by Clark (1971) and demonstrated by Skiba (1978), Majumdar and Mitra (1980) and Dechert and Nishimura (1983), the non-concavity of the Hamiltonian in the state variable plays a crucial role in generating unstable steady states and, possibly, a Dechert-Nishimura-Skiba (DNS) threshold.

The optimality condition suggests that along the growth-efficient drug-control path there is an equality between the marginal financial, private and social costs of controllers, $e^{-rt}[y + 2cL_g(t)]$, and the value to society of people prevented and rehabilitated by an additional *controller*, $-dl(t)$. By further considering the adjoint equation, $-l(t)$ — the present-value shadow cost of *users* — diminishes in a rate that is equal to the sum of the marginal diffusion of drugs and the ratio of the marginal return (MR) on *controllers* to the marginal costs (C') of *controllers*. That is,

$$\frac{\dot{l}(t)}{l(t)} = -F'(N(t);(1-b)L) - \overbrace{d(1-e)y}^{MR(L_g(t))} / \overbrace{(y + 2cL_g)}^{C'(L_g(t))} . \tag{6}$$

The evolution of the number of *controllers* along the growth-efficient path is given by the following no-arbitrage rule:

$$\dot{L}_g{}^o(t) = \frac{\overbrace{[r - F'(N(t);(1-b)L)]}^{UC(L_g{}^o)}\overbrace{[y + 2cL_g{}^o(t)]}^{C'(L_g{}^o)} - \overbrace{d(1-e)y}^{MR(L_g{}^o)}}{\underbrace{2c}_{C''}} . \tag{7}$$

The first term in the numerator of equation (7) is the instantaneously foregone gross national income stemming from an additional infinitesimal investment in control. It is equal to the product of the user cost (UC) of the control capital (namely, the fully productive people employed as *controllers*) and the financial and the private and social costs of employing an additional unit of control capital (a *controller*). The user-cost of the control capital (*controllers*) includes the social planner's rate of time preference (presumably the foregone national interest on any dollar spent on control) but is reduced by the instantaneous "infection" of non-using workers and *controllers* by an additional *user*, which is positive (negative) up to (beyond) the critical mass of

N^* *users*. The second term in the numerator of equation (7), $\boldsymbol{d}(1-\boldsymbol{e})y$, indicates the marginal return on control capital (*controllers*). The employment of a *controller* increases the number of the fully productive *nonusers* by $\boldsymbol{d}$. Thus, increases gross national income by $\boldsymbol{d}(1-\boldsymbol{e})y$.

The no-arbitrage rule suggests that the government's efficient employment of *controllers* changes during the planning horizon in accordance with the difference between the foregone gross national income stemming from, and the gross national income return on, an additional infinitesimal effort invested in drug control. If the loss of national income from employing an additional *controller* is greater (smaller) than the return on a *controller*, investment in additional control capital has to be postponed (brought forward). The inter-temporal change in the number of *controllers* is moderated by the coefficient ($2c$) of the associated marginal private and social cost. By adhering to this no-arbitrage rule the government facilitates the construction of a growth-efficient trajectory of the national portfolio of privately employed inputs comprising a fully effective labour force of $L-N(t)-L_g(t)$ *nonusers* and a less effective labour force of *N(t) users*.

4. Phase Portrait of Control and Use

The system comprising the no-arbitrage rule (equation (7)) and the net-loss of fully productive workers (equation (2)) has multiple steady states. By setting $\dot{L}_g = 0$ in equation (7), the steady-state levels of control effort are found from

$$L_g{}^{ss} = C''^{-1}(\boldsymbol{d}(1-\boldsymbol{e})y/[\boldsymbol{r} - F'(N_{ss};(1-\boldsymbol{b})L)]). \tag{8}$$

Recalling equation (2), the steady-state levels of control effort can also be found from

$$L_g^{ss} = F(N_{ss}; (1-b)L)/d .$$ (9)

Thus, the steady-state numbers of *users* satisfy the following equality

$$C''^{-1}(d(1-e)y/[r - F'(N_{ss}; (1-b)L)]) = F(N_{ss}; (1-b)L)/d .$$ (10)

The number and nature of these steady states is identified by the phase-plane analysis depicted in Figure 1 under the assumption that the planner's rate of time preference (r) is larger than the intrinsic diffusion rate of drugs ($F'(0; (1-b)L)$). A $r - F'(0; (1-b)L) > 0$ reflects a positive user cost of human capital in a drug-free society. It is also the rental cost of the control capital hired by the government to suppress an outburst of drug use.

The slope of the isocline $\dot{L}_g = 0$ is

$$\frac{dL_g}{dN}\bigg|_{\dot{L}_g=0} = \left[\frac{F''(N; (1-b)L)C'(t)}{[r - F'(N; (1-b)L)]C''(t)}\right] \overset{>}{\underset{<}{=}} 0 \quad \text{as} \quad F'(N; (1-b)L) \overset{>}{\underset{<}{=}} r, \text{ which in turn}$$

implies that $\dfrac{dL_g}{dN}\bigg|_{\dot{L}_g=0} \overset{>}{\underset{<}{=}} 0$ for $N \overset{<}{\underset{>}{=}} F''^{-1}(r)$ where $F''^{-1}(r) < N^*$. Since the government's rate of time preference (r) exceeds the intrinsic diffusion rate ($F'(0; (1-b)L)$), the isocline $\dot{L}_g = 0$ is negatively sloped in the entire phase plane.

As $\dfrac{d\dot{L}_g}{dN} = -F''C'/C'' > 0$, $\dot{L}_g$ is positive (negative) and depicted by upward (downward) pointed vertical arrows, in the region on the right (left)-hand-side of this isocline. The slope of the isocline $\dot{N} = 0$ is $\dfrac{dL_g}{dN}\bigg|_{\dot{N}=0} = \left[\dfrac{F'(N; (1-b)L)}{d}\right] \overset{>}{\underset{<}{=}} 0$ as

$N \overset{<}{\underset{>}{=}} N^*$ and hence this isocline is displayed by an inverted U-shaped curve in the phase plane. Since $\dfrac{d\dot{N}}{dL_g} = -d < 0$, $\dot{N}$ is negative (positive) and depicted by leftward

(rightward) pointed horizontal arrows, in the region above (below) this isocline. As displayed in Figure 1, the intersections of these isoclines define two unstable steady states — SS_1 and SS_2. As SS_1 (SS_2) is the steady state with a relatively high (low) control effort and hence a relatively low (high) number of users, it is not clear which of these steady states is DNI-superior: while SS_1 is supported by higher control cost, the drug-user intensive SS_2 is associated with a greater loss of labour production capacity.
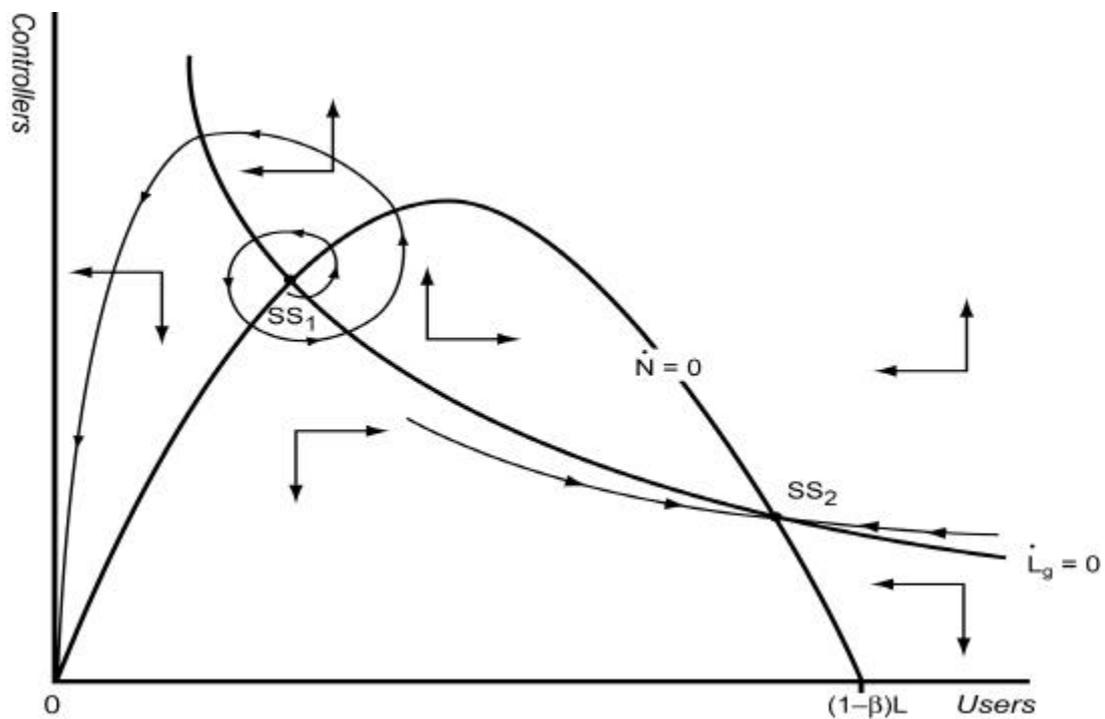

Figure 1: Phase-plane diagram with $r > F'(0;(1-b)L)$

The dynamic properties of SS_1 and SS_2 can be further assessed by considering the eigenvalues of the state-transition matrix (a) of the linearised system (in the vicinity of steady state) comprising the no-arbitrage equation (7) and the state equation (2)

12

$$m_{1,2} = 0.5\left\{ \boldsymbol{r} \pm \sqrt{\boldsymbol{r}^2 - 4\underbrace{\{[\boldsymbol{r} - F'(N_{ss})]F'(N_{ss}) - \boldsymbol{d}F''(N_{ss})C'_{ss}\}}_{\det a}} \right\}. \quad (11)$$

As $\boldsymbol{r} > F'(0)$, $F'' < 0$ and $C' > 0$ then $\boldsymbol{r} - F'(N_{SS}) > 0$ and $-\boldsymbol{d}F''(N_{ss})C'(t) > 0$ for both steady states.

In the case of SS_1, $F' > 0$ (since $N_{SS1} < N^*$) and hence the first term in $\det a$ is positive. In this case, $\det a > 0$ and hence SS_1 is not a saddle point. It is an unstable focus when $[\boldsymbol{d}F''(N_{SS1})C'_{SS1} - 4F'(N_{SS1})^2]/[\boldsymbol{r} - 4F'(N_{SS1})] < \boldsymbol{r} < 4F'(N_{SS1})$ and a node otherwise. In the case of SS_2, $F' < 0$ (since $N_{SS2} > N^*$) and hence the first term in $\det a$ is negative, rendering SS_2 a saddle point as long as $\boldsymbol{r} < F'(N_{SS2}) + \underbrace{\boldsymbol{d}C'_{SS2}F''(N_{SS2})/F'(N_{SS2})}_{\Omega > 0}$. For higher rates of time preference, $\det a(N_{SS2}, L_g{}^{SS2}) > 0$ implying that SS_2 is a node when $F'(N_{SS2}) + \Omega < \boldsymbol{r} < 2\sqrt{\det a(N_{SS2}, L_g{}^{SS2})}$ and an unstable focus for even higher discount rates.

It follows from the above arguments that the depiction of SS_1 and SS_2 in Figure 1 as an unstable focus point and a saddle point, respectively, which bears a resemblance to the phase portrait in the unconstrained model of Tragler *et al.* (2001), is possible in our framework if, in addition to exceeding $F'(0)$, the planner's discount rate satisfies the inequality

$$\frac{\boldsymbol{d}F''(N_{SS1})C'_{SS1} - 4F'(N_{SS1})^2}{\boldsymbol{r} - 4F'(N_{SS1})} < F'(N_{SS2}) + \boldsymbol{d}C'_{SS2}[F''(N_{SS2})/F'(N_{SS2})] < 4F'(N_{SS1})$$

or, equivalently (and observing that $\boldsymbol{r} - 4F'(N_{SS1}) < 0$), if

$$F'(0) < \boldsymbol{r} < 4F'(N_{SS1}) + \frac{\boldsymbol{d}F''(N_{SS1})C'_{SS1} - 4F'(N_{SS1})^2}{F'(N_{SS2}) + \boldsymbol{d}C'_{SS2}[F''(N_{SS2})/F'(N_{SS2})]}$$

where $F'(N_{SS2}) + dC'_{SS2}[F''(N_{SS2})/F'(N_{SS2})] < 4F'(N_{SS1})$. When these inequalities are satisfied, the low-control and large user number steady state SS_2 has a stable manifold consisting of two arms. One arm leads to SS_2 from North-West, indicating that it is optimal to gradually reduce control effort and allow the number of *users* to rise. The other arm converges to SS_2 from South-East revealing that it is optimal to gradually increase control effort and thereby reduce the number of *users*. The high-control steady state, SS_1, is an asymptotically unstable focus. That is, if initially the system is off that steady state, diverging oscillations in both the number of *users* and the number of *controllers* characterise the path maximising the sum of the discounted disposable national incomes. These oscillations display periods of high control effort leading eventually to a decline in number of *users*, followed by periods of reduced control effort subsequently leading to a rise in the number of *users*.

Similar to the special case described by Tragler *et al.* (2001) of unconstrained control budget and slowly diminishing returns on treatment, as proposed by Clark (1971, 1976) and proven by Majumdar and Mitra (1980) for the analogous case of efficient management of renewable natural resources with a convex-concave growth function in the state variable, there exists a critical point (DNS threshold) separating the basin of attraction around the high control and small user number equilibrium SS_1 from the convergent arm leading to the low control and large user number equilibrium SS_2. If the initial number of *users* is lower than the DNS threshold, it is optimal for the government to eliminate drug-use by applying, for a while, a concerted drug-control effort which subsequently diminishes as the number of *users* dwindles.[9] This

---

[9] It is unlikely in reality for illicit drug use to be completely eliminated because as user numbers dwindle, it becomes increasingly difficult for officials to identify remaining *users*. It is also difficult to envisage government employing the control effort require to eliminate a problem which is yet to become apparent.

policy is portrayed by the downward-sloping final segment of the unwinding spiral leading to the boundary equilibrium $N = 0 = L_g$.[10] However, if the initial number of *users* exceeds the DNS threshold, it is optimal for the government to gradually reduce effort and allow the number of *users* to rise and converge to the SS_2 level. There the cost of controlling any further increase in the number of *users* is relatively low due to their small net inflow (recalling that $F'' < 0$), which more than compensates for the loss of production capacity due to the number of *users* rising to N_{SS2}.

If the intrinsic diffusion rate of drug use exceeds the government's rate of time preference (i.e., $r < F'(0; (1 - b)L)$), the isocline $\dot{L}_g = 0$ has an inverted U-shape as displayed in Figure 2.
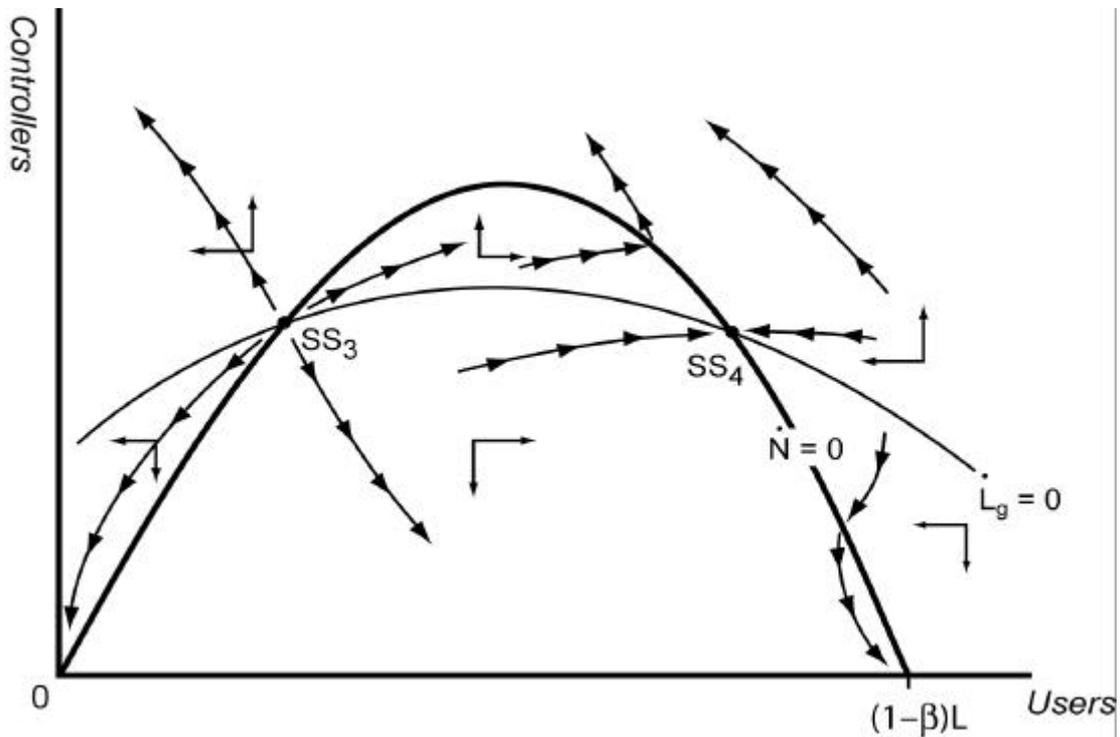

Figure 2: Phase-plane diagram with $r < F'(0; (1 - b)L)$

---

[10] Recalling equation (9) and that $F(0; (1 - b)L) = 0$ (see assumption 4), $L_g = 0$ when $N = 0$.

Recalling that $F'' < 0$ and $C' > 0$, the second term in $\det a$ is positive (namely, $-dF''(N_{SS})C'(t) > 0$) for both SS_3 and SS_4. In the case of SS_3, $N_{SS3} < N^*$ and hence $F'(N_{SS3}) > 0$, but the sign of $r - F'(N_{SS3})$ is not clear. In the case of SS_4, $N_{SS4} > N^*$ and hence $F'(N_{SS4}) < 0$, $r > F'(N_{SS4})$ and so the first term in $\det a$ is negative. While the nature of the steady states is not clearly revealed from the eigenvalues, the directions of the horizontal and vertical arrows (obtained, respectively, by differentiating $\dot{N}$ with respect to L_g and $\dot{L}_g$ with respect to N) indicate that also in this case the stationary points are asymptotically unstable: SS_3 is a node and SS_4 is, possibly, a saddle point. Elimination of drug use is possible and is a growth efficient policy for initial use-control combinations in some phases. In other phases it is efficient to let the number of *users* grow and converge to N_{SS4} or to the upper bound $(1 - b)L$.

5. Steady States and the Effects of the Model Parameters

To facilitate the analysis of the effects of the model parameters on the steady states, the diffusion function (F) is taken to be logistic

$$F(t) = aN(t)[1 - \frac{N(t)}{(1 - b)L}] \tag{12}$$

with $0 < a < 1$ indicating the intrinsic proliferation rate of drug use. In this case, the aforementioned no-arbitrage rule and the state-equation are displayed by the following system of non-linear differential equations:

$$\dot{L}_g(t) = \{\{r - a[1 - 2N(t)/(1 - b)L(t)]\}[(y + 2cL_g(t)] - d(1 - e)y\}/2c \tag{13}$$

$$\dot{N}(t) = aN(t)[1 - N(t)/(1 - b)L(t)] - dL_g(t). \tag{14}$$

By substituting $\dot{L}_g = 0 = \dot{N}$ into this system,

$$L_g{}^{ss} = \frac{1}{2c}\left[\frac{\boldsymbol{d}(1-\boldsymbol{e})y}{\boldsymbol{r}-\boldsymbol{a}+\dfrac{2\boldsymbol{a}N_{ss}}{(1-\boldsymbol{b})L}} - y\right] \tag{15}$$

where N_{ss} satisfies the following polynomial

$$\frac{2\boldsymbol{a}}{(1-\boldsymbol{b})L}N_{ss}{}^3 + (\boldsymbol{r}-3\boldsymbol{a})N_{ss}{}^2 - [(\boldsymbol{d}y/c)+(\boldsymbol{r}-\boldsymbol{a})(1-\boldsymbol{b})L]N_{ss} \tag{16}$$
$$= \boldsymbol{d}[\boldsymbol{r}-\boldsymbol{a}-\boldsymbol{d}(1-\boldsymbol{e})]y(1-\boldsymbol{b})L/2c\boldsymbol{a}$$

Consequently, the stationary level of gross national income (GNI_{ss}) is given by

$$GNI_{ss} = [L-(1-\boldsymbol{e})N_{ss}]y \tag{17}$$

while the stationary disposable national income (DNI_{ss}) is given by

$$DNI_{ss} = [L-(1-\boldsymbol{e})N_{ss}-L_g{}^{ss}]y - cL_g{}^{ss^2}. \tag{18}$$

A change in any of the model parameters affects the numbers of *users* and *controllers* and, consequently, the disposable national incomes in the steady states. From equation (18), DNI_{SS} decreases with both N_{SS} and $L_g{}^{ss}$ and is affected, through N_{SS} and $L_g{}^{ss}$, by any model parameter $\boldsymbol{g}$ with an elasticity that is a linear combination of the elasticities of the control and use with respect to that parameter:

$$\boldsymbol{x}_{DNI,\boldsymbol{g}} = -[(y+2cL_g{}^{ss})\boldsymbol{x}_{L_g,\boldsymbol{g}}L_g{}^{ss} + (1-\boldsymbol{e})y\boldsymbol{x}_{N,\boldsymbol{g}}N_{ss}]/DNI_{ss}. \tag{19}$$

The directions of the effects of parameter changes on the steady-states cannot be assessed by total differentiation of equations (15), (16) and (18) (see, for instance, $\partial N_{SS}/\partial\boldsymbol{g}$ in Appendix B). Nevertheless, insight on the effects of a parameter change on the steady-state levels of the endogenous variables can be gained from the shifts of the isoclines $\dot{L}_g = 0$ and $\dot{N} = 0$. Table 1 summarises the effect of an increase in each

parameter on the isoclines and thus on the high and low control steady-state levels of N, L_g and, in recalling equation (18), disposable national income for the case where $r - F'(0; (1-b)L) > 0$. The effects of all (some) of the model parameters on the low (high) control stationary levels of disposable national income are not clear due to opposing or unclear effects on the stationary numbers of *users* and *controllers*. The high-control stationary level of disposable national income level rises, due to a decline in the numbers of *users* and *controllers*, with the planner's discount rate, with the gradient of the private and social marginal cost of control effort and with the relative productivity of *users*; but declines, due to an increase in control spending and in the number of *users*, with the full-potential personal income.

Table 1: The effects on isoclines and steady states of increased parameter values

Parameters→	L	r	d	a	b	c	y	e
Isocline $\dot{L}_g = 0$	up	down	up	down	down	down	up	down
Isocline $\dot{N} = 0$	up	nil	down	up	down	nil	nil	nil
Lg^{ss_1}	rises	declines	?	?	declines	declines	rises	declines
N^{ss_1}	?	declines	rises	declines	?	declines	rises	declines
DNI_{SS1}	**?**	**rises**	**?**	**?**	**?**	**rises**	**declines**	**rises**
Lg^{ss_2}	rises	declines	rises	declines	declines	declines	rises	declines
N^{ss_2}	?	rises	declines	rises	?	rises	declines	rises
DNI_{SS2}	?	?	?	?	?	?	?	?

6. Extensions

Two extensions of the model are outlined in this section. The first extension distinguishes, as in Tragler et al. (2001), between two types of control activities—prevention and rehabilitation—and shows how their efficient levels of implementation depend upon their relative effectiveness and cost. The second extension incorporates societal costs of disharmony between *nonusers* and *users*.

6.1 Efficient Mix of Rehabilitation and Prevention

The government's effort in controlling the number of *users* may take different modes that for simplicity we categorise as rehabilitation and prevention. As indicated in the introduction, we define prevention as law enforcement and treatment activities that deter initiation, and rehabilitation as a process where long-term dependent users are restored to a drug-free existence. The optimal investment of effort in rehabilitation and prevention depends on the effectiveness and social cost differentials between these activities. The convexity of the private and social costs of prevention and rehabilitation ensures the optimality of mixed effort as long as those differentials are not too high in favour of one the activities.

Letting $0 \leq f(t) \leq 1$ be the proportion of $L_g(t)$ invested in rehabilitation at t, then the *user-nonuser* conversion equation (2) and the control cost equation (3) are modified as follows:

$$\dot{N}(t) = F(N(t);(1-b)L) - [d_r f(t) + d_p (1-f(t))]L_g(t) \tag{20}$$

$$C(t) = yL_g(t) + [c_r f(t)^2 + c_p (1-f(t))^2]L_g(t)^2 \tag{21}$$

where d_r and d_p are positive scalars denoting the marginal conversion and the marginal deterrent of the effort invested in rehabilitation and prevention, respectively; and where c_r and c_p are positive scalars denoting the coefficients of the private and social costs associated with investments of effort in rehabilitation and prevention, respectively.

It is possible that, due to its involuntary nature and its direct and external effects, prevention leads to higher levels of crime, violence and drug-quality uncertainty than does an equal investment of effort in rehabilitation. In this case, rehabilitation generates lower personal and social costs than prevention for any equal

level of effort—$c_r < c_p$. Due to the convexity of the private and social costs, however, there exists a sequence of portfolios of drug-use control activities with $0 < \boldsymbol{f}^*(t) < 1$ for every $t \in (0, \infty]$ that is growth-superior to investment in rehabilitation exclusively $(\boldsymbol{f}(t) = 1)$ even when, in addition to $c_r < c_p$, $\boldsymbol{d}_r \geq \boldsymbol{d}_p$, as long as c_p and $\boldsymbol{d}_r$ are not enormously larger than c_r and $\boldsymbol{d}_p$, respectively.[11]

As shown in Appendix C, the trajectory of the mixed effort that maximises the sum of the disposable national incomes accruing over an infinite planning horizon is given by the differential equation:

$$\dot{\boldsymbol{f}}(t) - [\boldsymbol{r} - F'(N(t);(1-\boldsymbol{b})L) - \dot{L}_g(t)/L_g(t)]\boldsymbol{f}(t)$$
$$= -\{[\boldsymbol{r} - F'(N(t);(1-\boldsymbol{b})L) - \dot{L}_g(t)/L_g(t)]c_p + 0.5(\boldsymbol{d}_r - \boldsymbol{d}_p)(1-\boldsymbol{e})y/L_g(t)\}/(c_r + c_p)$$
$$.(22)$$

In turn, the efficient rehabilitative effort share in steady state is

$$\boldsymbol{f}_{SS} = \frac{c_p}{c_r + c_p} - \frac{0.5(\boldsymbol{d}_p - \boldsymbol{d}_r)(1-\boldsymbol{e})y}{[\boldsymbol{r} - F'(N_{SS};(1-\boldsymbol{b})L)](c_r + c_p)L_g{}^{SS}}. \tag{23}$$

Since $L_g{}^{SS}$ and N_{SS} are endogenous, this expression alone can only be used for assessing the *direct* effects of the model parameters on $\boldsymbol{f}_{SS}$. In this respect, note the crucial roles of the difference between the discount rate and the stationary marginal drug-use diffusion and the difference between the marginal effectiveness of prevention and rehabilitation that lead to the following propositions.

If the discount rate is greater (smaller) than the stationary marginal drug-use diffusion then the direct effect of the prevention-rehabilitation marginal effectiveness differential $(\boldsymbol{d}_p - \boldsymbol{d}_r)$ on $\boldsymbol{f}_{SS}$ is negative (positive).

---

[11] See discussion of the relative effectiveness of prevention and rehabilitation in Gerstein, Johnson, Harwood, Fontain, Suter and Malloy (1994), Rydell, Caulkins and Everingham (1996), Caulkins, Rydell, Schwabe and Chiese (1997), and Crane, Rivolo and Comfort (1997).

If the discount rate is greater than the stationary marginal diffusion of drug-use then the direct effect of the rehabilitation's private and social cost coefficient c_r on $\boldsymbol{f}_{SS}$ is positive, zero, or negative when $\boldsymbol{d}_p - \boldsymbol{d}_r$ is greater than, equal to, or smaller than, $[\boldsymbol{r} - F'(N_{SS};(1-\boldsymbol{b})L)]L_g{}^{SS}c_p / 0.5(1-\boldsymbol{e})y$, respectively.

If $\boldsymbol{d}_p$ is smaller than $\boldsymbol{d}_r$ and the discount rate is smaller than the stationary marginal diffusion of drug-use then the direct effect of c_r on $\boldsymbol{f}_{SS}$ is positive.

If $\boldsymbol{d}_p$ is greater (smaller) than $\boldsymbol{d}_r$ then the direct effect of the planner's rate of time preference on the stationary rehabilitative effort share is positive (negative).

If $\boldsymbol{d}_p$ is greater than $\boldsymbol{d}_r$ and the discount rate is greater than the stationary marginal drug-use diffusion then the direct effect of the prevention's private and social cost coefficient c_p on $\boldsymbol{f}_{SS}$ is positive.

If either (neither) $\boldsymbol{d}_p$ is smaller than $\boldsymbol{d}_r$ or (nor) the discount rate is smaller than the stationary marginal drug-use diffusion then the direct effect of personal full capacity output y on $\boldsymbol{f}_{SS}$ is positive (negative).

If either (neither) $\boldsymbol{d}_p$ is smaller than $\boldsymbol{d}_r$ or (nor) the discount rate is smaller than the stationary marginal drug-use diffusion then the direct effect of the *users'* relative productivity $\boldsymbol{e}$ on $\boldsymbol{f}_{SS}$ is negative (positive).

If $\boldsymbol{d}_p$ is smaller (greater) than $\boldsymbol{d}_r$ and the stationary marginal drug-use diffusion increases with the size of the vulnerable labour force $(1-\boldsymbol{b})L$ (as is the case when the diffusion function is logistic) then the direct effect of L on $\boldsymbol{f}_{SS}$ is positive (negative) whereas the direct effect of the drug-use resistant population share $\boldsymbol{b}$ on $\boldsymbol{f}_{SS}$ is negative (positive).

Finally, if $\boldsymbol{d}_p = \boldsymbol{d}_r$ then $\boldsymbol{f}_{SS} = c_p / (c_r + c_p)$.

6.2 Societal Disharmony

If substance abuse is sufficiently widespread, tensions between *users* and *nonusers* might arise.[12] It is possible that the level of societal disharmony intensifies, and hence social costs increase, as the difference between the number of *nonusers* and *controllers* and the number of *users* diminishes. The model is extended to this case.

We assume that the relationship between costs of societal disharmony (*CSDH*) and the population share of *users* conforms to an inverted U-shaped curve

$$CSDH(t) = CSDH_{\max} - \boldsymbol{m}[(N(t)/L) - 0.5]^2 \tag{24}$$

where $CSDH_{\max}$ is the maximum societal cost of disharmony that accrues when the population shares of *users* and *nonusers* are equal, and $\boldsymbol{m}$ is a positive scalar reflecting the moderating effect of the quadratic distance from equal population shares on the cost of societal disharmony. This assumption implies, in conjunction with the assumptions made earlier, that the instantaneous DNI, now the gross national income net of the financial and social costs of prevention and the costs of societal disharmony, is given by

$$DNI(t) = [L - (1-\boldsymbol{e})N(t) - L_g(t)]y - cL_g(t)^2 - [CSDH_{\max} - \boldsymbol{m}((N(t)/L) - 0.5)^2]. \tag{25}$$

Consequently, the efficient number of *controllers* is now

$$\hat{L}_g = \arg\max \int_0^\infty e^{-rt} \{[L - (1-\boldsymbol{e})N(t) - L_g(t)]y - cL_g(t)^2 - [SDH_{\max} - \boldsymbol{m}((N(t)/L) - 0.5)^2]\}dt \tag{26}$$

subject to the motion equation (2).

---

[12] In the case of AIDS, there exists tension between infected and non-infected people and incidences of atrocities have been reported.

The no-arbitrage rule associated with this modification is

$$\dot{\hat{L}}_g(t) = \frac{[\boldsymbol{r} - F'(N(t);(1-\boldsymbol{b})L)][y + 2c\hat{L}_g(t)] - \boldsymbol{d}\{(1-\boldsymbol{e})y + (2\boldsymbol{m}/L)[(N(t)/L) - 0.5]\}}{2c}$$

.(27)

(See details in Appendix D.)

Since an extra infinitesimal effort in reducing the number of *users* does not

necessarily reduce the level of societal disharmony, $\dot{\hat{L}}_g \underset{>}{\overset{<}{=}} \dot{L}_g{}^o$ as $\dfrac{N(t)}{L} \underset{<}{\overset{>}{=}} 0.5$. If the

number of *users* initially exceeds the number of *nonusers*, a rise in the control effort

reduces the groups' size-differential and thereby intensifies social tension. In this case

the efficient increase in control effort is smaller than would be the case were societal

disharmony ignored. Conversely, if the number of *nonusers* initially exceeds the

number of *users*, a rise in control effort increases the groups' size-differential and

hence reduces societal tension. In this case, the efficient increase in control effort is

larger than would be the case were societal disharmony ignored.

7. Summary and Conclusions

In many countries illicit drug use is a serious problem that reduces the number of fully

productive workers and thereby aggregate output. This paper presented a dynamic

control model — a hybrid of an epidemiological diffusion process and a macro

economic objective — with a special reference to illicit drugs. The model is generic

and may also be applicable to other hazardous substances and activities.

The model divided the labour force into fully productive workers who do not

use drugs and only partially productive users, and assumed that the use of drugs is

contagious. In addition to foregone private output, costs are borne by government and

society with the provision of control effort. Efficient management of the nation's

portfolio of human resources was proposed as a path of drug-control effort that maximises the present value of the stream of disposable national incomes. The efficient level of control varies during the planning horizon in accordance with the difference between the foregone gross national income stemming from an additional infinitesimal effort invested in control and the return, in terms of gross national income, on an additional unit of effort invested in that activity. The inter-temporal change in control effort is moderated by the coefficient of the associated marginal personal and social costs. The foregone national income was taken as the product of the user cost of the typical *controller* and the marginal financial and social costs of the control effort. The user cost of a *controller* rises with the rate of time preference but is moderated by the instantaneous marginal "infection" of the labour force by *users*.

The possible steady states of the system comprising this rule and the proliferation equation were found to be asymptotically unstable. It was shown that if the discount rate is larger than the intrinsic diffusion rate and lower than a critical level, there exists a threshold number of users below which it is efficient, from an economic growth perspective, to eliminate drug use, and above which it is efficient for the economy to converge to a steady state with a low-control effort and a high number of *users*. If the intrinsic diffusion rate exceeds the planner's discount rate, elimination of drug use is growth efficient for some initial use-control combinations whereas allowing the number of users grow and converge to either the low-control steady-state number or to the upper bound is growth efficient for others.

The effects of the model parameters on the high-control and low-control steady-state numbers of *users* and *controllers*, and subsequently on the steady state disposable national income, were analysed for the case where the planner's discount rate exceeds the intrinsic rate of drug-use diffusion. The effects of each model

parameter on the level of disposable national income in the low-control steady state are indeterminate. In the high-control steady state, disposable national income increases, due to a decline in the numbers of *users* and *controllers*, with the discount rate, with the gradient of the private and social marginal cost of control effort and with the relative productivity of *users*, but declines, due to increases in control spending and the number of *users*, with full-potential personal income.

The model was then expanded to reflect on the composition of the control effort. There are cost and benefit differences between rehabilitation and prevention. As long as these differences are not too high in favour of one of the activities, convexity of the private and social costs of prevention and rehabilitation ensures the optimality of mixed effort. The propositions concerning the direct effects of the drug-use control's benefit and cost coefficients, the personal full capacity output and the *users*' relative productivity on the stationary rehabilitation-prevention mix reflected the crucial roles of the difference between the marginal effectiveness of prevention and rehabilitation and the difference between the discount rate and the stationary marginal diffusion of drug-use.

The model was also expanded to incorporate tension between *users* and *nonusers*. It was shown that if the number of *users* initially exceeds the number of *nonusers*, raising the control effort diminishes the groups' size-differential and, in turn, intensifies societal tension. In this case, the efficient rise in control effort is smaller than would be the case were societal tension ignored. Conversely, when initially the number of *nonusers* exceeds the number of *users*, raising the control effort increases the groups' size-differential and subsequently reduces the level of societal tension. In this case, the efficient rise in control effort is larger than would be the case were societal disharmony ignored.

REFERENCES

Behrens, D.A., Caulkins, J.P., Tragler, G., and Feichtinger, G. (1997) "Controlling U.S. cocaine epidemic: prevention from light vs. treatment of heavy use". Working Paper 214, Department of Operations Research and Systems Theory, Vienna University of Technology.

Caulkins, J. P., Rydell, C. P., Schwabe, W. L. and Chiesa, J. (1997) *Mandatory Minimum Drug Sentences: Throwing Away the Key of the Taxpayers' Money?* RAND, Santa Monica, CA.

Chabet, J. (2002) Mexico's War on Drugs: No Margin for Maneuver. *Annals of the American Academy of Political and Social Science* 582,134-148.

Clark, C.W. (1971) "Economically optimal policies for the utilization of biologically renewable resources", *Mathematical Bioscience* 17, 245-268.

Clark, C.W. (1976) *Mathematical Bioeconomics*, New York: John Wiley & Sons.

Collins, D and Lapsley, H. (1991) *Estimating the Economic Costs of Drug Abuse in Australia*, National Campaign Against Drug Abuse, Monograph Series No. 15, Australian Government Printing Service.

Crane, B.D., Rivolo, A.R. and Comfort G.C. (1997) *An Empirical Examination of Counterdrug Interdiction Program Effectiveness*, Alexandria, Virginia: Institute for Defense Analysis.

Dechert, W.D., and Nishimura, K. (1983) "A Complete Characterization of Optimal Growth Paths in an Aggregated Model with a Non-Concave Production Function", *Journal of Economic Theory* 31(2), 332-354.

Ferrence, R. (2001) "Diffusion Theory and Drug Use", *Addiction* 96, 165-173.

Gerstein, D.R., Johnson, R.A., Hardwood, H.J., Fontain, D., Suter, N. and Malloy, K. (1994) *Evaluating Recovery Services: The California Drug and Alcohol Treatment Assessment*, National Opinion Researh Center, Chicago IL and Lewin-VHI faifax, VA.

Harwood, H., Fountain, D. and Livermore, G. (1998) *The Economic Costs of Alcohol and Drug Abuse in the United States-1992,* National Institutes of Health, Rockville, MD.

Kamien, M. and Schwartz, N. (1991) *Dynamic Optimization*, Second Edition, Amsterdam: North-Holland.

Kennally, G. (2001) Regulating the Trade in Recreational Drugs. *European Journal of Law and Economics* 11, no.1, 69-82.

Klein, A. (1999) "Nigeria and the Drugs War", *Review of African Political Economy* 26, no.79, 51-73.

Majumdar, M. and Mirta, T. (1980) "On Optimal Exploitation of a Renewable Resources in a Non-Convex Environment and the Minimum Safe Standard of Conservation", Working Paper No. 223, Cornell University.

Rydell, C.P., Caulkins, J. P., and Everingham, S. S. (1996) "Enforcement or Treatment: Modeling the Relative Efficacy of Alternatives for Controlling Cocaine", *Operations Research* 44(6), 687-695.

Skiba, A.K. (1978) "Optimal Growth with a Convex-Concave Production Function", *Econometrica* 46.3, 527-539.

Tragler, G., Caulkins, J.P., and Feichtinger, G. (2001), "Optimal Dynamic Allocation of Treatment and Enforcement in Illicit Drug Control", *Operations Research* 49.3, 352-362.

Thoumi, F. (2002) Illegal Drugs in Colombia: From Illegal Economic Boom to Social Crisis. *Annals of the American Academy of Political and Social Science* 582, 102-116.

Xie, X., Rehrn, J., Single, E., Robson, L., and Paul J. (1998) "The Economic Costs of Illicit Drug Use in Ontario", 1992. *Health Economics* 7, no.1, 81-85.

Appendix A: The necessary and sufficient conditions and the no-arbitrage rule

The (present value) Hamiltonian associated with equation (5) and equation (2) is

$$H(t) = e^{-rt}\{[L-(1-\boldsymbol{e})N(t)-L_g(t)]y - cL_g(t)^2\} + \boldsymbol{l}(t)[F(N(t);(1-\boldsymbol{b})L) - \boldsymbol{d}L_g(t)].$$
(A1)

The necessary conditions for maximum are:

$$\dot{\boldsymbol{l}}(t) = -\frac{\partial H(t)}{\partial N(t)} = e^{-rt}[(1-\boldsymbol{e})y] - \boldsymbol{l}(t)F'(N(t);(1-\boldsymbol{b})L)$$
(A2)

$$\frac{\partial H(t)}{\partial L_g(t)} = -e^{-rt}[y + 2cL_g(t)] - \boldsymbol{dl}(t) = 0$$
(A3)

equation (2) and the transversality condition $\lim_{t\to\infty} \boldsymbol{l}(t)N(t) = 0$.

Equation (6) is obtained by dividing both sides of equation (A2) by $\boldsymbol{l}$ and considering that by virtue of equation (A3) $\boldsymbol{l}(t) = -e^{rt}[y + 2cL_g(t)]/\boldsymbol{d}$. The no-arbitrage rule, equation (7), is obtained by differentiating the optimality condition (A3) with respect to t (singular control), substituting the information contained in conditions (A2) and (A3) for $\dot{\boldsymbol{l}}$ and $\boldsymbol{l}$, respectively, multiplying both sides by $e^{rt}/2c$ and rearranging terms. It can also be obtained by using Euler equation.

Appendix B: The effects of changes in the model parameters

The effects of changes in L, r, d, a, b, c, y, and e on N_{ss} are:

$$\frac{dN_{ss}}{dL} = \frac{\dfrac{2aN_{ss}^3}{(1-b)L^2} + (r-a)(1-b)N_{ss} + \dfrac{d[r-a-d(1-e)]y(1-b)}{2ca}}{\dfrac{6aN_{ss}^2}{(1-b)L} + 2(r-3a)N_{ss} - (dy/c) - (r-a)(1-b)L} \qquad (B1)$$

$$\frac{dN_{ss}}{dr} = \frac{(1-b)L[N_{ss} + dy/2ca] - N_{ss}^2}{\dfrac{6aN_{ss}^2}{(1-b)L} + 2(r-3a)N_{ss} - (dy/c) - (r-a)(1-b)L} \qquad (B2)$$

$$\frac{dN_{ss}}{dd} = \frac{[yN_{ss} + (r-a-2d)(1-b)L/2a]/c}{\dfrac{6aN_{ss}^2}{(1-b)L} + 2(r-3a)N_{ss} - (dy/c) - (r-a)(1-b)L} \qquad (B3)$$

$$\frac{dN_{ss}}{da} = \frac{3N_{ss}^2 - \{[2N_{ss}^3/(1-b)L] + (1-b)L[N_{ss} + d(1-b)L(ya - ((r-a-d(1-e))y)/2ca^2]\}}{\dfrac{6aN_{ss}^2}{(1-b)L} + 2(r-3a)N_{ss} - (dy/c) - (r-a)(1-b)L}$$

$$(B4)$$

$$\frac{dN_{ss}}{db} = \frac{-\left(\dfrac{2aN_{ss}^3}{(1-b)^2L} + (r-a)LN_{ss} + \dfrac{dL\{[r-a-d(1-e)]y\}}{2ca}\right)}{\dfrac{6aN_{ss}^2}{(1-b)L} + 2(r-3a)N_{ss} - (dy/c) - (r-a)(1-b)L} \qquad (B5)$$

$$\frac{dN_{ss}}{dc} = \frac{-d[yN_{ss} + (r-a-d(1-e))y(1-b)L/2a]/c^2}{\dfrac{6aN_{ss}^2}{(1-b)L} + 2(r-3a)N_{ss} - (dy/c) - (r-a)(1-b)L} \qquad (B6)$$

$$\frac{dN_{ss}}{dy} = \frac{(d/c)\{N_{ss} + [r-a-d(1-e)](1-b)L/2a\}}{\dfrac{6aN_{ss}^2}{(1-b)L} + 2(r-3a)N_{ss} - (dy/c) - (r-a)(1-b)L} \qquad (B7)$$

$$\frac{dN_{ss}}{d(1-\boldsymbol{e})} = \frac{-\boldsymbol{d}^2 y(1-\boldsymbol{b})L/2c\boldsymbol{a}}{\dfrac{6\boldsymbol{a}N_{ss}^2}{(1-\boldsymbol{b})L} + 2(\boldsymbol{r}-3\boldsymbol{a})N_{ss} - (\boldsymbol{d}y/c) - (\boldsymbol{r}-\boldsymbol{a})(1-\boldsymbol{b})L}. \tag{B8}$$

30

Appendix C: The no–arbitrage rule of equation (22)

The decision problem in section 6 is

$$\max_{\{\boldsymbol{f},L_g\}} \int_0^\infty e^{-rt}\{[L-(1-\boldsymbol{e})N(t)-L_g(t)]y-[c_r\boldsymbol{f}(t)^2+c_p(1-\boldsymbol{f}(t))^2]L_g(t)^2\}dt \qquad (C1)$$

subject to the state equation (20) and $N(\infty)\geq 0$. The Hamiltonian associated with this problem is

$$H(t)=e^{-rt}\{[L-(1-\boldsymbol{e})N(t)-L_g(t)]y-[c_r\boldsymbol{f}(t)^2+c_p(1-\boldsymbol{f}(t))^2]L_g(t)^2\}$$
$$+\boldsymbol{l}(t)\{F(N(t),(1-\boldsymbol{b})L)-[\boldsymbol{d}_r\boldsymbol{f}(t)+\boldsymbol{d}_p(1-\boldsymbol{f}(t))]L_g(t)\}$$

$$\qquad (C2)$$

The necessary conditions for maximum are

$$\dot{\boldsymbol{l}}(t)=-\frac{\partial H(t)}{\partial N(t)}=e^{-rt}(1-\boldsymbol{e})y-\boldsymbol{l}(t)F'(N(t);(1-\boldsymbol{b})L) \qquad (C3)$$

$$\frac{\partial H(t)}{\partial \boldsymbol{f}}=-2e^{-rt}[(c_r+c_p)\boldsymbol{f}(t)-c_p]L_g(t)^2-\boldsymbol{l}(t)(\boldsymbol{d}_r-\boldsymbol{d}_p)L_g(t)=0 \qquad (C4)$$

$$\frac{\partial H(t)}{\partial L_g(t)}=-e^{-rt}\{y+2[c_r\boldsymbol{f}^2+c_p(1-\boldsymbol{f})^2]L_g\}-\boldsymbol{l}[\boldsymbol{d}_r\boldsymbol{f}+\boldsymbol{d}_p(1-\boldsymbol{f})]=0. \qquad (C5)$$

Equation (22) is obtained by dividing both side of equation C4 by L_g, differentiating with respect to t, substituting the right-hand-side of equation C3 for $\dot{\boldsymbol{l}}$ and the expression of $\boldsymbol{l}$ from equation C4, and rearranging terms.

Appendix D: The no-arbitrage rule of equation (27)

The Hamiltonian associated with this decision problem is

$$H(t) = e^{-rt}\{[L-(1-\boldsymbol{e})N(t)]y - C(L_g(t)) - [CSDH_{max} - \boldsymbol{m}((N(t)/L)-0.5)^2]\} \\ + \boldsymbol{l}(t)[F(N(t);(1-\boldsymbol{b})L) - \boldsymbol{d}L_g(t)]$$.(D1)

The necessary conditions for maximum are:

$$\dot{\boldsymbol{l}}(t) = -\frac{\partial H(t)}{\partial N(t)} = e^{-rt}\{(1-\boldsymbol{e})y + (2\boldsymbol{m}/L)[((N(t))/L)-0.5]\} \\ - \boldsymbol{l}(t)F'(N(t);(1-\boldsymbol{b})L)$$ (D2)

$$\frac{\partial H(t)}{\partial L_g(t)} = -e^{-rt}C'(L_g(t)) - \boldsymbol{d}\boldsymbol{l}(t) = 0$$ (D3)

equation (2) and the transversality condition $\lim_{t\to\infty} \boldsymbol{l}(t)N(t) = 0$.

The no-arbitrage rule, equation (27), is obtained by differentiating the optimality condition (D3) with respect to t (singular control), substituting the information contained in conditions (D2) and (D3) for $\dot{\boldsymbol{l}}$ and $\boldsymbol{l}$, respectively, multiplying both sides by $e^{rt}/C''(L_g(t))$ and rearranging terms.