Macroeconomic Dynamics, **10**, 2006, 145–164. Printed in the United States of America.
DOI: 10.1017.S1365100506050140

ARTICLES

MACROECONOMIC ASPECTS OF SUBSTANCE ABUSE: DIFFUSION, PRODUCTIVITY AND OPTIMAL CONTROL

AMNON LEVY
City College of New York, City University of New York
and
University of Wollongong

FRANK NERI
University of Wollongong

DIETER GRASS
Vienna University of Technology

This paper deals with macroeconomic aspects of widespread substance abuse with a reference to illicit drugs as an example. Substance abuse impedes the productivity of the labor force and reduces economic growth. Workers are either nonusers and therefore fully productive, a number of whom are employed by the government in drug-control activities, or users who are only partially productive. Efficient management of the nation's portfolio of workers involves eradicating drug use when initial user numbers are lower than a critical level, but allows user numbers to rise to, and be accommodated at, a stationary level when initial user numbers exceed a critical level.

Keywords: Substance Abuse, Labor Productivity, National Income, Optimal Control

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 labor, 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 delegitimating the regime, has diminished trust and increased transaction costs, violence, impunity, security costs, and has promoted highly speculative investments and capital

Address correspondence to: Professor Amnon Levy, Economics Discipline, School of Economics and Information Systems, University of Wollongong, NSW 2522, Australia; e-mail: amnon_levy@uow.edu.au.

© 2006 Cambridge University Press 1365-1005/06 $12.00

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

flight (Thoumi, 2002). Brito and Intriligator (1992) portray the drug lords as Stackelberg leaders who achieve their preferred outcome by appropriate transfers to the government or the guerrillas. Other notable examples are Yemen, where addiction, although 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 also has had deleterious effects in Mexico (Chabet, 2002) and Nigeria (Klein, 1999). Illicit drugs also have adversely affected rich countries. For example, Harwood, Fountain, and Livermore (1998) estimate the cost of drug abuse and dependence in United States 1992 at U.S. $98 billion. 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 Canadian $489 million. The estimate of Collins and Lapsley (1991) for Australia in 1988 is around Australian $1.6 billion.

Although almost all drugs once were 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 the 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-minimizing mix of spending on enforcement and treatment.

This paper, although 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 on aggregate production, economic growth, and the debilitating effects of illicit drugs. More specifically, we focus on the relationships between disposable national income, the decomposition of the labor force between users and nonusers, and government control efforts. We construct a macrodynamic optimization model to describe rational public investment of effort in controlling the spread of illicit drugs and their adverse effect on the labor force and economic growth. Drug control effort is determined so as to maximize the sum of the discounted disposable national incomes in the presence of fully productive nonusing workers and less productive drug users. Drug use is modeled 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 this paper also can be used for analyzing the growth-efficient control on other potentially harmful substances and activities such as alcohol, tobacco, junk-food products, and HIV-infectious behavior. Therefore, although we refer explicitly to illicit drugs, our macrodynamic approach is sufficiently general for dealing conceptually with the abuse of any potentially debilitating substance or activity.[1]

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

The paper continues as follows. Section 2 introduces the building blocks of our macrodynamic drug-control model in which the nation's labor force is divided into nonusing fully productive workers and less productive drug users. Section 3 presents and interprets the rule for drug-control effort that maximizes the present value of the stream of the net national incomes stemming from the country's using and nonusing workers. Two interior equilibriums are identified in Section 4: 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 analyzed 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 analyzed 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, in which 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 become 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 macrodynamic model focuses on the relationship between the use of illicit drugs, labor productivity, aggregate income, and government drug-control effort. For tractability, we ignore capital, capital accumulation, and technological changes and assume that labor is the sole factor of production. More specifically, our model is based on the following assumptions.

Assumption 1 (labor-force size and composition): The size of the working-age population is time invariant and equal to L,[2] of which $0 \leq N(t) \leq (1 - \beta)L$ people use illicit drugs, hereafter *users*, where $0 \leq \beta < 1$ denotes the share of the population absolutely unsusceptible to illicit drugs.

Assumption 2 (drugs and employment): L_g members of the labor force are drug-control officers, hereafter *controllers*, employed by the government and providing a composite service of prevention and rehabilitation.[3] The remaining $L - L_g$ members are employed in the private sector. *Controllers* 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 $L - N(t) - L_g(t)$

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

nonusers is y then the instantaneous output of a *user* is:

$$y_n = \varepsilon y, \quad 0 \leq \varepsilon < 1, \tag{1}$$

where y is a positive time-invariant scalar, and ε is the relative productivity of a *user* with $\varepsilon = 0$ indicating total incapacitation and $0 < \varepsilon < 1$ partial incapacitation.[4]

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[5] drug-control function (δL_g):

$$\dot{N}(t) = F(N(t); (1-\beta)L) - \delta L_g(t), \tag{2}$$

where δ 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 becomes more socially acceptable.[6] Because of 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 $0 < N < (1-\beta)L$. 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 $N^* < (1-\beta)L$. Thereafter, $F' < 0$ reflecting a negative marginal diffusion, which is a result of a dominant natural attrition effect, and leading to $F = 0$ when $N = (1-\beta)L$. Furthermore, in order to retain a nonnegative state variable, we assume that $F \leq 0$ for all $N \leq 0$. Recalling that $F > 0$ for all $0 < N < (1-\beta)L$, then $F(0; (1-\beta)L) = 0$.[7] Recalling also that $\delta L_g \geq 0$, we merely require that the terminal value of N is nonnegative ($N(\infty) \geq 0$) to ensure the nonnegativity of $N(t)$ for every t in the planning horizon.[8] Nevertheless, the case $N = 0$ has to be considered explicitly (see Appendix B).

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 government 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:

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

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

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* also are identical but proportionally lower, a progressive tax scheme would subsidise *users* at the expense of *nonusers* and would affect the supply of labor. In contrast, a lump-sum tax treats *users* and *nonusers* equally and does not affect the supply of labor. Concerns about intra- and intergenerational diffusion of substance abuse and its debilitating effect on personal and aggregate production motivates nonusers to bear part of the costs of the government control effort.

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:

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

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 intertemporally assessing national economic benefits. Maximizing the sum of the discounted instantaneous DNIs is equivalent, in this case, to maximizing 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 maximizes 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,

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

subject to the motion equation (2) and $N(\infty) \geq 0$, and where ρ 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. Because the co-state variable $\lambda(t)$ multiplying $[F(N(t); (1 - \beta)L) - \delta L_g(t)]$ is nonpositive and the diffusion function $F(N(t); (1 - \beta)L)$ is assumed to be concave in N, the Hamiltonian is nonconcave in N and Mangasarian's theorem on the sufficiency of 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 nonconcavity of the Hamiltonian in the

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

state variable plays a crucial role in generating unstable steady states and, possibly, a Dechert-Nishimura-Skiba (DNS) point.

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*, $e^{-\rho t}[y + 2cL_g(t)]$, and the value to society of people prevented and rehabilitated by an additional *controller*, $-\delta\lambda(t)$. By further considering the adjoint equation, $-\lambda(t)$—the present-value shadow cost of *users*—diminishes at 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,

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = -F'(N(t); (1-\beta)L) - \overbrace{\delta(1-\varepsilon)y}^{MR(L_g(t))} \Big/ \overbrace{(y + 2cL_g)}^{C'(L_g(t))}. \tag{6}$$

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

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

The first term in the numerator of (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, private, and social costs of employing an additional unit of control capital (a c*ontroller*). 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 nonusing 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 (7), $\delta(1-\varepsilon)y$, indicates the marginal return on control capital (*controllers*). The employment of a *controller* increases the number of fully productive *nonusers* by δ and hence increases gross national income by $\delta(1-\varepsilon)y$.

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 intertemporal 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

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

construction of a growth-efficient trajectory of the national portfolio of privately employed inputs comprising a fully effective labor force of $L - N(t) - L_g(t)$ *nonusers* and a less effective labor force of $N(t)$ *users*.

4. PHASE PORTRAIT OF CONTROL AND USE

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

$$L_g^{ss} = \frac{\delta(1-\varepsilon)y}{2c[\rho - F'(N_{ss}; (1-\beta)L)]} - y/2c. \tag{8}$$

Recalling (2), the steady-state levels of control effort also satisfy:

$$L_g^{ss} = F(N_{ss}; (1-\beta)L)/\delta. \tag{9}$$

In turn, the steady-state numbers of *users* satisfy the following equality:

$$\frac{\delta(1-\varepsilon)y}{2c[\rho - F'(N_{ss}; (1-\beta)L)]} - y/2c = F(N_{ss}; (1-\beta)L)/\delta. \tag{10}$$

These steady-state levels of control and use are found at the intersections of the isoclines $\dot{L}_g = 0$ and $\dot{N} = 0$ in the phase-plane diagram. As displayed in Figure 1, the isocline $\dot{L}_g = 0$ is negatively sloped in the entire phase plane. This is because of $\frac{dL_g}{dN|\dot{L}_g=0} = \frac{\delta(1-\varepsilon)yF''(N;(1-\beta)L)}{[\rho - F'(N;(1-\beta)L)]^2}$ and $F''(N; (1-\beta)L) < 0$. Furthermore, as $\frac{d\dot{L}_g}{dN} = -F''C'/C'' > 0$, $\dot{L}_g$ is positive (negative) and depicted by upward (downward) pointed vertical arrows, in the region to the right (left) of this isocline. The slope of the isocline $\dot{N} = 0$ is

$$\frac{dL_g}{dN_{|\dot{N}=0}} = \left[\frac{F'(N; (1-\beta)L)}{\delta}\right] \gtreqless 0$$

as $N \gtreqless N^*$ and hence this isocline is displayed by an inverted U-shaped curve. Because $\frac{d\dot{N}}{dL_g} = -\delta < 0$, $\dot{N}$ is negative (positive) and depicted by leftward (rightward) pointed horizontal arrows, in the region above (below) this isocline. The intersections of these isoclines and the directions of the horizontal and vertical arrows in their vicinity define two unstable steady states (in the positive quadrant)—an unstable focus SS_1 and a saddle point 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. Although SS_1 is supported by higher control cost, the drug-user intensive SS_2 is associated with a greater loss of labor production capacity. As a marginal solution, SS_0 at the origin also has to be considered.

The depiction of SS_1 and SS_2 as an unstable focus point and a saddle point, respectively, bears a resemblance to the phase portrait in the unconstrained model

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

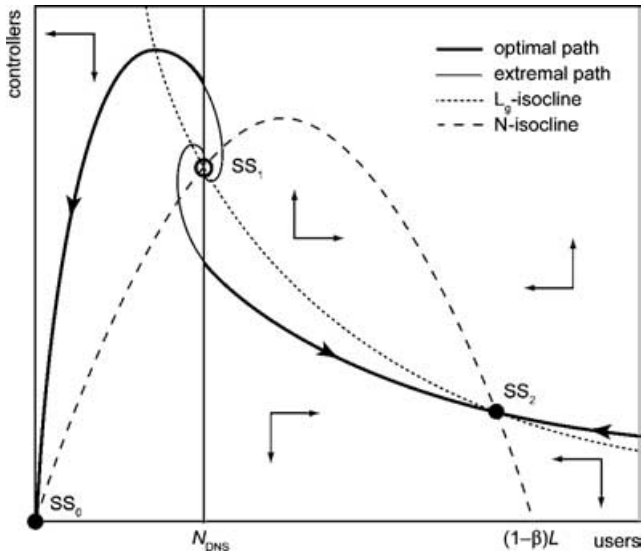


FIGURE 1. Phase-plane diagram.

of Tragler et al. (2001). 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*. By contrast, the path leading from North-East to SS_0 also has to be considered as an optimal solution. This implies that it is optimal to increase the number of *controllers* at first and, falling below some number of *users*, the number of *controllers* can be reduced. From this kind of phase portrait, an unstable focus between two optimal long-run steady states, the existence of a DNS point can directly be concluded (see Feichtinger and Hartl 1986).[9] To calculate the DNS point explicitly, one has to compare the objective functions along the paths leading to the long-run steady states. The number of *users* where the objective values coincide is a DNS point (N_{DNS} in Figure 1). At this point, the policy maker is indifferent between increasing or decreasing control efforts, as both of them are equally optimal. If the initial number of *users* is smaller than N_{DNS}, it is optimal for the government to eradicate drug use by applying, for a while, a concerted drug-control effort, which subsequently diminishes as the number of *users* dwindles.[10] This policy is portrayed by the downward-sloping final segment of the unwinding spiral leading to the boundary equilibrium $N = 0 = L_g$ (see details in Appendix B). However, if the initial number of *users* is larger than N_{DNS}, it is optimal for the government to gradually reduce effort and allow the number of *users* to rise and converge to the level at SS_2. There, the cost of controlling any further increase in the number of *users* is relatively low because of their small net inflow (recalling

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

that $F'' < 0$), which more than compensates for the loss of production capacity because of the number of *users* rising to N_{SS_2}.

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:

$$F(t) = \alpha N(t)\left[1 - \frac{N(t)}{(1-\beta)L}\right], \tag{11}$$

with $0 < \alpha < 1$ 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 nonlinear differential equations:

$$\dot{L}_g(t) = \{\{\rho - \alpha[1 - 2N(t)/((1-\beta)L(t))]\}[y + 2cL_g(t)] - \delta(1-\varepsilon)y\}/2c \tag{12}$$

$$\dot{N}(t) = \alpha N(t)[1 - N(t)/((1-\beta)L(t))] - \delta L_g(t). \tag{13}$$

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

$$L_g^{ss} = \frac{1}{2c}\left[\frac{\delta(1-\varepsilon)y}{\rho - \alpha + \frac{2\alpha N_{ss}}{(1-\beta)L}} - y\right], \tag{14}$$

where N_{ss} satisfies the following polynomial:

$$\frac{2\alpha}{(1-\beta)L}N_{ss}^3 + (\rho - 3\alpha)N_{ss}^2 - [(\delta y/c) + (\rho - \alpha)(1-\beta)L]N_{ss}$$
$$= \delta[\rho - \alpha - \delta(1-\varepsilon)]y(1-\beta)L/2c\alpha. \tag{15}$$

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

$$GNI_{ss} = [L - (1-\varepsilon)N_{ss}]y, \tag{16}$$

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

$$DNI_{ss} = \left[L - (1-\varepsilon)N_{ss} - L_g^{ss}\right]y - cL_g^{ss^2}. \tag{17}$$

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 (17), 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 γ with an elasticity that is a linear combination of the elasticities of control and use with respect to that parameter:

$$\xi_{DNI,\gamma} = -\left[\left(y + 2cL_g^{ss}\right)\xi_{L_g,\gamma}L_g^{ss} + (1-\varepsilon)y\xi_{N,\gamma}N_{ss}\right]/DNI_{ss}. \tag{18}$$

The directions of the effects of parameter changes on the steady states cannot be assessed by total differentiation of (14), (15), and (17) (see, for instance, $\partial N_{SS}/\partial\gamma$

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

TABLE 1. The effects on isoclines and steady states of increased parameter values

Parameters→	L	ρ	δ	α	β	c	y	ε
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_{SS_1}	**?**	**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_{SS_2}	**?**	**?**	**?**	**?**	**?**	**?**	**?**	**?**

in Appendix C). Nevertheless, insights 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 summarizes 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 (17), disposable national income. The effects of all (some) of the model parameters on the low- (high-) control stationary levels of disposable national income are not clear because of opposing or unclear effects on the stationary numbers of *users* and *controllers*. The high-control stationary level of disposable national income rises, because of 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, because of an increase in control spending and in the number of *users*, with the full-potential personal income.

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 on 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 categorize 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 in which 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

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

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 favor of one of the activities.

Letting $0 \leq \phi(t) \leq 1$ 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:

$$\dot{N}(t) = F(N(t); (1-\beta)L) - [\delta_r \phi(t) + \delta_p(1-\phi(t))]L_g(t) \tag{19}$$

$$C(t) = yL_g(t) + [c_r \phi(t)^2 + c_p(1-\phi(t))^2]L_g(t)^2, \tag{20}$$

where δ_r and δ_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, because of 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 $c_r < c_p$. Because of the convexity of the private and social costs, however, there exists a sequence of portfolios of drug-use control activities with $0 < \phi^*(t) < 1$ for every $t \in (0, \infty]$ that is growth-superior to investment in rehabilitation exclusively ($\phi(t) = 1$) even when, in addition to $c_r < c_p$, $\delta_r \geq \delta_p$, as long as c_p and δ_r are not enormously larger than c_r and δ_p, respectively.[11]

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

$$\begin{aligned}\dot{\phi}(t) - [\rho - F'(N(t); (1-\beta)L) - \dot{L}_g(t)/L_g(t)]\phi(t) \\ = -\{[\rho - F'(N(t); (1-\beta)L) - \dot{L}_g(t)/L_g(t)]c_p \\ + 0.5(\delta_r - \delta_p)(1-\varepsilon)y/L_g(t)\}/(c_r + c_p).\end{aligned} \tag{21}$$

In turn, the efficient rehabilitation effort share in steady state is:

$$\phi_{SS} = \frac{c_p}{c_r + c_p} - \frac{0.5(\delta_p - \delta_r)(1-\varepsilon)y}{[\rho - F'(N_{SS}; (1-\beta)L)](c_r + c_p)L_g^{SS}}. \tag{22}$$

Because 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 ϕ_{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.

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

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 ($\delta_p - \delta_r$) on ϕ_{SS} is negative (positive).

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 ϕ_{SS} is positive, zero, or negative when $\delta_p - \delta_r$ is greater than, equal to, or smaller than, $[\rho - F'(N_{SS}; (1-\beta)L)]L_g^{SS} c_p / 0.5(1-\varepsilon)y$, respectively.

If δ_p is smaller than δ_r and the discount rate is smaller than the stationary marginal diffusion of drug-use, then the direct effect of c_r on ϕ_{SS} is positive.

If δ_p is greater (smaller) than δ_r then the direct effect of the planner's rate of time preference on the stationary share of rehabilitation effort is positive (negative).

If δ_p is greater than δ_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 ϕ_{SS} is positive.

If either (neither) δ_p is smaller than δ_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 ϕ_{SS} is positive (negative).

If either (neither) δ_p is smaller than δ_r or (nor) the discount rate is smaller than the stationary marginal drug-use diffusion then the direct effect of the *users*' relative productivity ε on ϕ_{SS} is negative (positive).

If δ_p is smaller (greater) than δ_r and the stationary marginal drug-use diffusion increases with the size of the vulnerable labor force $(1-\beta)L$ (as is the case when the diffusion function is logistic), then the direct effect of L on ϕ_{SS} is positive (negative), whereas the direct effect of the drug-use resistant population share β on ϕ_{SS} is negative (positive).

Finally, if $\delta_p = \delta_r$ then $\phi_{SS} = c_p/(c_r + c_p)$.

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.

$$CSDH(t) = CSDH_{\max} - \mu[(N(t)/L) - 0.5]^2 \qquad (23)$$

$CSDH_{\max}$ is the maximum societal cost of disharmony that accrues when the population shares of *users* and *nonusers* are equal, and μ 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

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

societal disharmony, is given by:

$$DNI(t) = [L - (1-\varepsilon)N(t) - L_g(t)]y - cL_g(t)^2 - [CSDH_{\max} - \mu((N(t)/L) - 0.5)^2]. \quad (24)$$

Consequently, the efficient number of *controllers* is now:

$$\hat{L}_g = \arg\max \int_0^\infty e^{-\rho t}\{[L - (1-\varepsilon)N(t) - L_g(t)]y - cL_g(t)^2 - [SDH_{\max} - \mu((N(t)/L) - 0.5)^2]\}\, dt, \quad (25)$$

subject to the motion equation (2). The no-arbitrage rule associated with this modification is (see details in Appendix E):

$$\dot{\hat{L}}_g(t) = \frac{[\rho - F'(N(t);(1-\beta)L)][y + 2c\hat{L}_g(t)] - \delta\{(1-\varepsilon)y + (2\mu/L)[(N(t)/L) - 0.5]\}}{2c}. \quad (26)$$

Because an extra infinitesimal effort in reducing the number of *users* does not necessarily reduce the level of societal disharmony, $\dot{\hat{L}}_g \lessgtr \dot{L}_g^o$ as $N(t)/L \lessgtr 0.5$. 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 presents 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 also may be applicable to other hazardous substances, to epidemic control, and to socially undesirable activities such as crime.

The model divides the labor force into fully productive workers who do not use drugs and only partially productive users, and assumes 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. Concerns about intra- and intergenerational diffusion of substance abuse motivates nonusers to bear

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

part of the costs of the government control effort. Efficient management of the nation's portfolio of human resources is proposed as a path of drug-control effort that maximizes 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 intertemporal change in control effort is moderated by the coefficient of the associated marginal personal and social costs. The foregone national income is 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 labor force by *users*.

The steady states of the system comprising the derived equation of the efficient change in the control effort and the assumed proliferation equation are found to be an unstable focus with a high level of control and a low number of users and a saddle point with a low level of control and a high number of users. Below a critical number of users it is optimal, from an economic growth perspective, to eradicate drug use by applying a concerted initial control effort and then by gradually reducing control. Above that critical number of users it is optimal for the government to apply a much lower level of initial control and to gradually reduce control so that the economy converges to the steady state with low-control effort and a high number of *users*.

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, also were analyzed. 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, because of 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, because of increases in control spending and the number of *users*, with full-potential personal income.

The model was expanded to reflect on the composition of control effort. There are cost and benefit differences between rehabilitation and prevention. As long as these differences are not too high in favor 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 drug 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 also was expanded to incorporate tension between *users* and *nonusers*. It was shown that if the number of *users* initially exceeds the number of

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

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.

NOTES

1. Although the analysis refers to substance abuse, the framework developed is also applicable to schooling (where users are high school drop-outs and controllers are teachers), crime (where users are criminals and controllers are police officers) and public health management (where users are sick people and controllers are physicians).

2. A time-invariant population size is also assumed 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.

3. 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.

4. In an intertemporal 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: $\dot{\varepsilon} > 0$ as a *user* becomes gradually accustomed to the adverse effects of drugs and develops some level of tolerance to those effects. However, in an intertemporal analysis of *national* production and substance abuse, such as the present one, $1 - \varepsilon$ denotes the average loss of production capacity within the group of *users*. Because of 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 a result of attrition—death of heavy and old *users*—and also rehabilitation and maturation of others. Because the membership in the group of *users* changes continually it is difficult to justify $\dot{\varepsilon} > 0$ or $\dot{\varepsilon} < 0$. Moreover, because of 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.

5. An alternative concave specification—$R(L_g)$, $R' > 0$, $R'' < 0$—reflecting diminishing marginal control requires $-\lambda(t)R''(Lg(t)) < 2c$ for an interior solution to the maximization problem described in Section 3 to exist, where c is the private and social cost coefficient indicated in assumption 5 and λ is the co-state variable of the maximization problem's Hamiltonian. It also is 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.

6. 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 (2001) for a review.

7. Note that the logistic diffusion function used in Section 4, $F(t) = \alpha N(t)[1 - N(t)/((1 - \beta)L)]$ with $0 < \alpha < 1$ denoting the intrinsic diffusion rate $N^* = 0.5(1 - \beta)L$, satisfies these assumptions.

8. 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 nonnegativity requirement. Thus, our terminal nonnegativity restriction assures nonnegativity of N throughout.

9. The existence of a DNS point in the present framework is similar to the one in the special case described by Tragler et al. (2001) of unconstrained control budget and slowly diminishing returns on

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

treatment, and to the critical point in the analogous case of efficient management of renewable natural resources with a convex-concave growth function in the state variable proposed by Clark (1971, 1976) and proven by Majumdar and Mitra (1980).

10. 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 required for eliminating a problem that is yet to become apparent.

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

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

REFERENCES

Behrens, Doris A., Jonathon P. Caulkins, Gernot Tragler, and Gustav Feichtinger (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.

Brito, Dagobert L. and Michael D. Intriligator (1992) Narco-traffic and guerrilla warfare: a new symbiosis. *Defence Economics* 3(4), 263–274.

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

Chabet, Jorge (2002) Mexico's war on drugs: no margin for manoeuvre. *Annals of the American Academy of Political and Social Science* 582, 134–148.

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

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

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

Crane, Barry D., A. Rex Rivolo, and Gary C. Comfort (1997) *An Empirical Examination of Counter-Drug Interdiction Program Effectiveness*. Alexandria, VA: Institute for Defence Analysis.

Dechert, W. Davis and Kazuo Nishimura (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, Roberta (2001) Diffusion theory and drug use. *Addiction* 96, 165–173.

Feichtinger, Gustav and Richard F. Hartl (1986) *Optimale Kontrolle Ökonomischer Prozesse*. New York: Walter de Gruyter.

Gerstein, Dean R., Robert A. Johnson, Henrick J. Harwood, Douglas Fountain, Natalie Suter, and Kay Malloy (1994) *Evaluating Recovery Services: The California Drug and Alcohol Treatment Assessment*. Chicago IL and Lewin-VHI Fairfax, VA: National Opinion Research Center.

Harwood, Henrick J., Douglas Fountain, and Gina Livermore (1998) *The Economic Costs of Alcohol and Drug Abuse in the United States—1992*. Rockville, MD: National Institutes of Health.

Kamien, Morton I. and Nancy L. Schwartz (1991) *Dynamic Optimization*, 2nd Ed. Amsterdam: North-Holland.

Kennally, Gerald (2001) Regulating the trade in recreational drugs. *European Journal of Law and Economics* 11(1), 69–82.

Klein, Axel (1999) Nigeria and the drugs war. *Review of African Political Economy* 26(79), 51–73.

Majumdar, Makul and Tapan Mitra (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.

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

Rydell, C. Peter, Jonathon P. Caulkins, and Susan S. Everingham (1996) Enforcement or treatment: modelling 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, Gernot, Jonathon P. Caulkins, and Gustav Feichtinger (2001) Optimal dynamic allocation of treatment and enforcement in illicit drug control. *Operations Research* 49(3), 352–362.

Thoumi, Francisco (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, Xiaodi, Jurgen Rehrn, Eric Single, Lynda Robson, and Josh Paul (1998) The economic costs of illicit drug use in Ontario: 1992. *Health Economics* 7(1), 81–85.

APPENDIX A

THE NECESSARY CONDITIONS AND THE NO-ARBITRAGE RULE

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

$$H(t) = e^{-\rho t}\{[L - (1-\varepsilon)N(t) - L_g(t)]y - cL_g(t)^2\} + \lambda(t)[F(N(t); (1-\beta)L) - \delta L_g(t)]. \quad \textbf{(A1)}$$

The necessary conditions for maximum are:

$$\dot{\lambda}(t) = -\frac{\partial H(t)}{\partial N(t)} = e^{-\rho t}[(1-\varepsilon)y] - \lambda(t)F'(N(t); (1-\beta)L) \quad \textbf{(A2)}$$

$$\frac{\partial H(t)}{\partial L_g(t)} = -e^{-\rho t}[y + 2cL_g(t)] - \delta\lambda(t) = 0, \quad \textbf{(A3)}$$

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

(6) is obtained by dividing both sides of (A2) by λ and considering that by virtue of (A3) $\lambda(t) = -e^{\rho t}[y + 2cL_g(t)]/\delta$. The no-arbitrage rule (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 $\dot{\lambda}$ and λ, respectively, multiplying both sides by $e^{\rho t}/2c$ and rearranging terms. It also can be obtained by using Euler equation.

APPENDIX B

THE NECESSARY CONDITIONS FOR THE STATE CONSTRAINT $N(t) \geq 0$

Considering (A1), the Lagrangian $\Lambda(t)$ for the state constraint $N(t) \geq 0$ becomes

$$\Lambda(t) = H(t) + \nu(t)N(t). \quad \textbf{(B1)}$$

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

Following Kamien and Schwartz (1991, p. 230), the necessary optimality conditions are:

$$\Lambda_{L_g}(t) = -e^{-\rho t} y - 2cL_g(t) - \delta\lambda(t) = 0 \tag{B2}$$

$$\dot{\lambda}(t) = -\frac{\partial \Lambda(t)}{\partial N(t)} = e^{-\rho t}(1-\varepsilon)y - \lambda(t)F'(N(t); (1-\beta)L) - \nu(t) \tag{B3}$$

$$\nu(t) \geq 0 \tag{B4}$$

$$N(t)\nu(t) = 0. \tag{B5}$$

Therefore, the constrained canonical system is:

$$\dot{N}(t) = -\delta L_g(t) \tag{B6}$$

$$\dot{\lambda}(t) = e^{-\rho t}(1-\varepsilon)y - \lambda(t)F'(0; (1-\beta)L) - \nu(t). \tag{B7}$$

Furthermore, as the Legendre-Clebsh condition $\Lambda_{L_g L_g}(t) = -2c < 0$ is satisfied it is assured that the costate $\lambda(t)$ and control $L_g(t)$ are continuous at switching points where the state constraint becomes active (see, for example, Feichtinger and Hartl, 1986). From (B6) it can be seen that the origin in the state control space is a candidate for an optimal solution. To prove (B4) we get from (B2), $L_g = 0$ and because of the continuity of the costate variable at the switching time t_0:

$$\lambda(t) = -e^{-\rho t} y/\delta \quad \text{for } t \geq t_0. \tag{B8}$$

Setting now $\dot{\lambda}(t) = 0$ and substituting (B8) the following equality for $\nu(t)$ with $t \geq t_0$ holds.

$$\nu(t) = e^{-\rho t} y \left(1 - \varepsilon + \frac{F'(0; (1-\beta)L)}{\delta}\right). \tag{B9}$$

As $\varepsilon < 1$, $F'(N(t); (1-\beta)L) > 0$, $y > 0$ and $\delta > 0$ have been assumed, this proves $\nu(t) > 0$ for $t \geq t_0$. Hence, $N = 0 = L_g$ satisfies the necessary optimality conditions.

APPENDIX C

THE EFFECTS OF CHANGES IN THE MODEL PARAMETERS

The effects of changes in L, ρ, δ, α, β, c, y, and ε on N_{ss} are as follows:

$$\frac{dN_{ss}}{dL} = \frac{\dfrac{2\alpha N_{ss}^3}{(1-\beta)L^2} + (\rho-\alpha)(1-\beta)N_{ss} + \dfrac{\delta[\rho-\alpha-\delta(1-\varepsilon)]y(1-\beta)}{2c\alpha}}{\dfrac{6\alpha N_{ss}^2}{(1-\beta)L} + 2(\rho-3\alpha)N_{ss} - (\delta y/c) - (\rho-\alpha)(1-\beta)L} \tag{C1}$$

$$\frac{dN_{ss}}{d\rho} = \frac{(1-\beta)L[N_{ss} + \delta y/2c\alpha] - N_{ss}^2}{\dfrac{6\alpha N_{ss}^2}{(1-\beta)L} + 2(\rho-3\alpha)N_{ss} - (\delta y/c) - (\rho-\alpha)(1-\beta)L} \tag{C2}$$

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

$$\frac{dN_{ss}}{d\delta} = \frac{[yN_{ss} + (\rho - \alpha - 2\delta)(1-\beta)L/2\alpha]/c}{\dfrac{6\alpha N_{ss}^2}{(1-\beta)L} + 2(\rho - 3\alpha)N_{ss} - (\delta y/c) - (\rho - \alpha)(1-\beta)L} \tag{C3}$$

$$\frac{dN_{ss}}{d\alpha} = \frac{3N_{ss}^2 - \{[2N_{ss}^3/(1-\beta)L] + (1-\beta)L\left[N_{ss} + \delta(1-\beta)L(y\alpha - (\rho - \alpha - \delta(1-\varepsilon))y)/2c\alpha^2\right]\}}{\dfrac{6\alpha N_{ss}^2}{(1-\beta)L} + 2(\rho - 3\alpha)N_{ss} - (\delta y/c) - (\rho - \alpha)(1-\beta)L} \tag{C4}$$

$$\frac{dN_{ss}}{d\beta} = \frac{-\left(\dfrac{2\alpha N_{ss}^3}{(1-\beta)^2 L} + (\rho - \alpha)LN_{ss} + \dfrac{\delta L\{[\rho - \alpha - \delta(1-\varepsilon)]y\}}{2c\alpha}\right)}{\dfrac{6\alpha N_{ss}^2}{(1-\beta)L} + 2(\rho - 3\alpha)N_{ss} - (\delta y/c) - (\rho - \alpha)(1-\beta)L} \tag{C5}$$

$$\frac{dN_{ss}}{dc} = \frac{-\delta[yN_{ss} + (\rho - \alpha - \delta(1-\varepsilon))y(1-\beta)L/2\alpha]/c^2}{\dfrac{6\alpha N_{ss}^2}{(1-\beta)L} + 2(\rho - 3\alpha)N_{ss} - (\delta y/c) - (\rho - \alpha)(1-\beta)L} \tag{C6}$$

$$\frac{dN_{ss}}{dy} = \frac{(\delta/c)\{N_{ss} + [\rho - \alpha - \delta(1-\varepsilon)](1-\beta)L/2\alpha\}}{\dfrac{6\alpha N_{ss}^2}{(1-\beta)L} + 2(\rho - 3\alpha)N_{ss} - (\delta y/c) - (\rho - \alpha)(1-\beta)L} \tag{C7}$$

$$\frac{dN_{ss}}{d(1-\varepsilon)} = \frac{-\delta^2 y(1-\beta)L/2c\alpha}{\dfrac{6\alpha N_{ss}^2}{(1-\beta)L} + 2(\rho - 3\alpha)N_{ss} - (\delta y/c) - (\rho - \alpha)(1-\beta)L}. \tag{C8}$$

APPENDIX D

THE NO-ARBITRAGE RULE OF EQUATION (21)

The decision problem in Section 6 is:

$$\max_{\{\phi, L_g\}} \int_0^\infty e^{-\rho t}\{[L - (1-\varepsilon)N(t) - L_g(t)]y - [c_r\phi(t)^2 + c_p(1-\phi(t))^2]L_g(t)^2\}\,dt \tag{D1}$$

subject to the state equation (19) and $N(\infty) \geq 0$. The Hamiltonian associated with this problem is

$$\begin{aligned} H(t) = {} & e^{-\rho t}\{[L - (1-\varepsilon)N(t) - L_g(t)]y - [c_r\phi(t)^2 + c_p(1-\phi(t))^2]L_g(t)^2\} \\ & + \lambda(t)\{F(N(t), (1-\beta)L) - [\delta_r\phi(t) + \delta_p(1-\phi(t))]L_g(t)\} \end{aligned} \tag{D2}$$

The necessary conditions for maximum are

$$\dot{\lambda}(t) = -\frac{\partial H(t)}{\partial N(t)} = e^{-\rho t}(1-\varepsilon)y - \lambda(t)F'(N(t); (1-\beta)L) \tag{D3}$$

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press

$$\frac{\partial H(t)}{\partial \phi} = -2e^{-\rho t}[(c_r + c_p)\phi(t) - c_p]L_g(t)^2 - \lambda(t)(\delta_r - \delta_p)L_g(t) = 0 \qquad \textbf{(D4)}$$

$$\frac{\partial H(t)}{\partial L_g(t)} = -e^{-\rho t}\{y + 2[c_r\phi^2 + c_p(1-\phi)^2]L_g\} - \lambda[\delta_r\phi + \delta_p(1-\phi)] = 0. \qquad \textbf{(D5)}$$

(21) is obtained by dividing both side of (D4) by L_g, differentiating with respect to t, substituting the right-hand side of (D3) for $\dot{\lambda}$ and the expression of λ from (D4), and rearranging terms.

APPENDIX E

THE NO-ARBITRAGE RULE OF EQUATION (26)

The Hamiltonian associated with this decision problem is

$$H(t) = e^{-\rho t}\{[L - (1-\varepsilon)N(t)]y - C(L_g(t)) - [CSDH_{\max} - \mu((N(t)/L) - 0.5)^2]\}$$
$$+ \lambda(t)[F(N(t); (1-\beta)L) - \delta L_g(t)]. \qquad \textbf{(E1)}$$

The necessary conditions for maximum are

$$\dot{\lambda}(t) = -\frac{\partial H(t)}{\partial N(t)} = e^{-\rho t}\{(1-\varepsilon)y + (2\mu/L)[((N(t))/L) - 0.5]\} - \lambda(t)F'(N(t); (1-\beta)L), \qquad \textbf{(E2)}$$

$$\frac{\partial H(t)}{\partial L_g(t)} = -e^{-\rho t}C'(L_g(t)) - \delta\lambda(t) = 0, \qquad \textbf{(E3)}$$

(2) and the transversality condition $\lim_{t\to\infty} \lambda(t)N(t) = 0$. The no-arbitrage rule (26) is obtained by differentiating the optimality condition (E3) with respect to t (singular control), substituting the information contained in conditions (E2) and (E3) for $\dot{\lambda}$ and λ, respectively, multiplying both sides by $e^{\rho t}/C''(L_g(t))$ and rearranging terms.

https://doi.org/10.1017/S1365100506050140 Published online by Cambridge University Press