Skiba Thresholds in Optimal Control of Illicit Drug Use*

Gustav Feichtinger[†] Gernot Tragler[‡]

April 15, 2002

Abstract

During the last quarter of a century, dynamic models with multiple steady states have been studied in numerous areas of economics. The corresponding history-dependence of optimal paths constitutes a low-level form of complexity. The purpose of the present paper is to illustrate this fact for three selected models in the control of illicit drug consumption, which are validated with empirical data. In the first model, the dynamics of the current U.S. cocaine epidemic subject to law enforcement and treatment is studied. The second part augments the first model by taking into account the fact that enforcement activities influence not only the drug dynamics but also property crime. The third model investigates the influence of methadone maintenance treatment on the spread of HIV/HCV among injection drug users. In all three cases, a positive feedback effect, i.e. state-dependent initiation, is responsible for the occurrence of 'Skiba points' separating the basins of attraction of the multiple steady states.

1 Introduction

In his seminal paper 'Crime and punishment: an economic approach', the later Nobel prize winner in economics, Gary S. Becker, poses the following provoking question:

*This research was partly financed by the Austrian National Bank (ÖNB) under grant No. 9414. We thank Maria Dworak and Julia Balta for providing files from their Master theses that were necessary to produce Sections 3 and 4, respectively. Finally, we thank Christian Almeder for revision of a preliminary version of this paper.

†Vienna University of Technology, Institute for Econometrics, Operations Research and Systems Theory; Argentinierstr. 8/119, A-1040 Wien, Austria; Email: or@eos.tuwien.ac.at; Fax: +43 1 58801 11999; Phone: +43 1 58801 11927.

‡Vienna University of Technology, Institute for Econometrics, Operations Research and Systems Theory; Argentinierstr. 8/119, A-1040 Wien, Austria; Email: tragler@eos.tuwien.ac.at; Fax: +43 1 58801 11999; Phone: +43 1 58801 11920.

'How many offenses should be permitted, and how many offenders should go unpunished?'

(Becker, 1968, p. 170). Although Becker was not the first providing an economic approach for illegal behaviour and its punishment, his work was path-breaking, creating a whole new subject in economic theory, i.e. the economics of crime. According to his own words, Becker's main purpose was to answer the question, how many resources and how much punishment should be used to enforce different kinds of legislation. Given the cost of arresting and convicting offenders, and the damage from offenses, one should find those expenditures of resources and punishment that minimize the total social loss.

Becker's analysis was static, and surprisingly few followers took into consideration that most offending and law enforcement evolve over time, i.e. are essentially a dynamic process. Thus, a more realistic approach to combat crime has to include intertemporal aspects. In particular, the following dynamic extension of Becker's static supply function of offenses has been proposed by Caulkins (1993) in the context of a macro-dynamic description of the movement of dealers into and out of local drug markets under police enforcement. Comparing the dealers with firms and illicit drug markets with industries, where free entry and exit ensure zero long-run profit, Caulkins proposed that the rate of change of offenders depends on the expected utility from illegal activity compared with that from legal work. To model such a framework it is assumed that the potential criminals become offenders as soon as their individual utility expected from committing a crime exceeds the (average) income from an alternative, but legal activity. If their utility is smaller than the reservation wage, criminals will lower or even stop the number of offenses. In addition to that, the dynamics of offenders is reduced by the rate of apprehended criminals.

Using such an offenders' dynamics, Feichtinger et al. (1997) minimize the total discounted stream of social losses (as described above). By applying optimal control theory they are able to prove an interesting 'threshold behaviour' of optimal law enforcement policies. In particular, this means that there exists a critical level for offenses, denoted by N_c, in the following sense. If the initial number of offenders $N(0)$ is *above* N_c, then there is a long-run 'high' equilibrium (i.e. a long-run steady state which is a saddle point) which is gradually approached along the stable manifolds (both from below and above as long as $N(0)$ is greater than the threshold N_c). In economic terms, this means that the intertemporal trade-off between the damage from offenses and the law enforcement and punishment costs yields an upper (interior) equilibrium. This result answers the question 'how many offenses should be permitted' posed by Becker (1968) and mentioned above. In addition to Becker, the dynamic analysis provides the optimal path of law enforcement. It turns out that its structure makes economic sense.

However, if $N(0)$ is *below* the critical level N_c, then it is optimal to eradicate crime, i.e. it pays to enforce until the illegal market collapses. The steady-state equilibrium is at the (lower) boundary, and the law enforcement expenditures increase first, but finally decrease. For $N(0) = N_c$, the optimal control is non-unique, i.e. the optimal policy is discontinuous. For details see Feichtinger et al. (1997).

In most cases, it is the convexity of the Hamiltonian with respect to the state variable(s) implying this interesting result. However, it has been recently stressed, that the economically important threshold property is compatible with strict concavity (Feichtinger and Wirl, 2000). To come back to Feichtinger et al.'s (1997) model of optimal law enforcement, we might ask for the reason of the 'threshold behaviour' in that particular example. Since the dependence of the optimal paths on the initial condition is due to the existence of *multiple equilibria*, the question shifts to this property. Examining the model it is easily seen that the property is due to a special non-linearity originating in the dependence of the conviction probability on the total law enforcement expenditures (E) *per offender* (N). Thus, the *macro* level, N, influences the *individual* conviction probability p ($p = p\left(\frac{E}{N}\right)$).

This idea is of general importance in socio-economics. There are virtually dozens of examples in which the macro level influences the micro characteristics. Actually, in the literature those effects are well-known. Some decades ago, Schelling (1978) has already discussed 'micromotives and macrobehaviour'. It turns out that such dependencies of micro characteristics on the macro environment create *non-linearities* being rich enough to generate complex solution structures.

Let us now focus more specifically on the macro-micro impact in the economics of crime. In his excellent survey on the economics of corruption, Andvig (1991) assumes that the utility an individual receives from a given action depends on the choices of others in that individual's reference group. For instance, in an environment where corruption is the norm it would not be rational to 'stay clean'. The general idea is conveyed through what Andvig calls a *Schelling diagram* (see Schelling, 1973, p. 388) and is simple: the expected profitability of engaging in offending depends on the number of other people who do it, i.e. on the size of the reference group. Consider a binary choice, i.e. to offend and not to offend. The abscissa of the Schelling diagram measures the percentage of criminality in the reference group. The origin means that the society is completely 'clean', whereas an abscissa value of one says that the reference group is totally criminal. According to Schelling's idea, the individual utility function for offending as well as for non-offending depends in a characteristic manner on the society's offending behaviour. Without going into details (see the description given in Andvig, 1991, pp. 69-75), we may state that the utility for non-offending is *above* the profitability for offending for a clean society. On the other hand, a 'white sheep' in a herd of black sheeps has a comparable disadvantage. This configuration implies that the two utility functions intersect (at least) once. Assuming for simplicity that there is only one point of intersection B, then there are three equlilibria: two boundary equilibria A and C, where all are clean or all are criminal, respectively, and one interior equilibrium B. While A and C are stable, B is an *unstable* equlibrium point. At A all are non-offending and will prefer to stay that way, because their utility levels are above that of any offender. At C, the opposite is true with a similar conclusion. At B, however, any agent is indifferent between the offending and non-offending activity, but if only one more individual is offending, it will pay to become an offender. If only one person less is offending, (s)he will choose to be 'clean', too. Thus B is an unstable equilibrium. Note that this threshold behaviour of the unstable point B may be interpreted as some sort of prelude to a *DNS* point.[1]

This paper is organized as follows. In Section 2, a dynamic cost-benefit analysis for the current U.S. cocaine epidemic is discussed. Although it contains only one state, i.e. the number of illicit drug consumers, it illustrates a policy-oriented optimal control model for an empirical data set. In particular, the dynamic interaction of treatment spending and law enforcement expenditures is studied. For this model, there exist three long-run steady states, i.e. two stable equilibria (one interior and a boundary one) and one unstable fixed point. The Skiba threshold separating the two basins of attraction is calculated numerically. Section 3 extends this model by including property crime. Again a Skiba point is identified, where the economic interpretation is analogous to the model discussed in Section 2. Section 4 presents another optimal control model describing the influence of methadone maintenance treatment (MMT) on the spread of blood-borne diseases like HIV or HCV among injection drug users. The aim of the model is to find the optimal policy to minimize the discounted stream of overall costs arising from MMT and the social costs caused by new infections from HIV/HCV. The structure of the optimal paths contains a *DNS* threshold and makes economic sense. The concluding Section 5 delivers a comparative analysis of the existing control policies separated by *DNS* thresholds.

2 Law Enforcement and Treatment in a Model of the U.S. Cocaine Epidemic

Tragler et al. (2001) provide a one-state optimal control model of cocaine use in the U.S. with two controls. (For details see also Tragler, 1998). The state variable $A(t)$ describes the number of users at time t, while the controls are treatment spending $u(t)$ and law enforcement spending $v(t)$, respectively. The objective is

[1] In this paper, we will denote the thresholds separating the basins of attraction both as 'Skiba points/thresholds' and as '*DNS* points/thresholds', where the abbreviation '*DNS*' stands for Dechert-Nishimura-Skiba. While Skiba (1978) discussed such critical values, his proof was incomplete; see Dechert and Nishimura (1983) for a first sound proof.

to minimize the discounted sum of the costs associated with drug use plus the costs of drug control over some given time horizon.

In Tragler et al.'s (2001) model, law enforcement is directed against dealers, not users. More precisely, the focus is on enforcement's ability to act like a tax that drives up the cost of distributing drugs (Reuter and Kleiman, 1986) and consequently the drug price. As pointed out by Rydell and Everingham (1994), 'the money spent on supply control causes increases in the cost to producers of supplying the cocaine. That increased cost of supply gets passed along to the consumer as price increases, which in turn causes the number of users to decline as inflows to cocaine use decrease and outflows increase.' In other words, price-raising law enforcement has both a direct and an indirect effect: the direct effect of an increased price is expressed in a reduction of the amounts consumed, while the indirect effect comes from the fact that high prices suppress initiation and encourage desistance.

The model of enforcement's effect on price $p(t)$ is taken from Caulkins et al. (1997):

$$p(t) = p\left(A(t), v(t)\right) = d + e\frac{v(t)}{A(t) + \epsilon}, \tag{1}$$

where ϵ is an arbitrarily small constant that avoids division by zero. The parameter d captures the fact that prohibition itself forces suppliers to operate in inefficient ways (what Reuter, 1983, calls 'structural consequences of product illegality'). Because of enforcement swamping (cf. Kleiman, 1993), the marginal effectiveness of enforcement (e) is multiplied by enforcement effort relative to market size ($\frac{v(t)}{A(t)}$), not total enforcement effort.

Note that in what follows we may omit the time argument t to increase readability.

Consumption is modeled as $\theta A p^{-\omega}$, which is consistent with a constant elasticity model of per capita demand. The state equation has terms for initiation (I), outflow due to treatment (O_u), and the background rate of desistance (O_v) (cf. Fig. 1). The rate of initiation is an increasing function of the current number of users (kA^{α}) modulated by price:

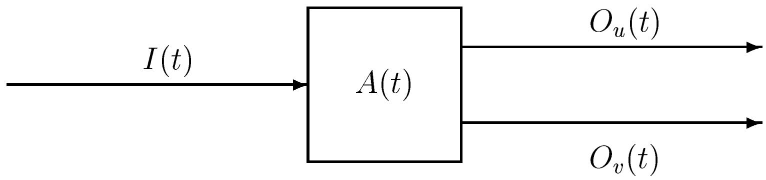
$$I(t) = kp\left(A(t), v(t)\right)^{-a} A(t)^{\alpha}. \tag{2}$$

The per capita rate of desistance is assumed to be a constant (μ) modulated by price:

$$O_v(t) = \mu p\left(A(t), v(t)\right)^{b} A(t). \tag{3}$$

Hence, in the absence of controls, the elasticity of the steady state number of users with respect to price is $-a-b$. The overall, or long-run, elasticity of demand is the sum of the elasticity of demand per capita and the price elasticity of the number of users, i.e. $-\left(a+b+\omega\right)$.

Figure 1: Flow diagram for the model by Tragler et al. (2001).



Outflow due to treatment is modeled as being proportional to treatment spending per capita raised to an exponent z that reflects diminishing returns,[2] with a small constant in the denominator (δ) that prevents division by zero. In particular,

$$O_u(t) = c\left(\frac{u(t)}{A(t)+\delta}\right)^z A(t),$$

where c denotes the treatment efficiency proportionality constant.

Summarizing, the state equation is given by

$$\dot{A}(t) = kp\left(A(t), v(t)\right)^{-a} A(t)^{\alpha} - c\left(\frac{u(t)}{A(t)+\delta}\right)^z A(t) - \mu p\left(A(t), v(t)\right)^b A(t)$$

with the objective specified as

$$\min_{u(t)\geq 0, v(t)\geq 0} \int_0^\infty e^{-rt}\left(\kappa\theta A(t) p\left(A(t), v(t)\right)^{-\omega} + u(t) + v(t)\right) dt,$$

where r denotes the discount rate and κ is the social cost per unit of consumption.

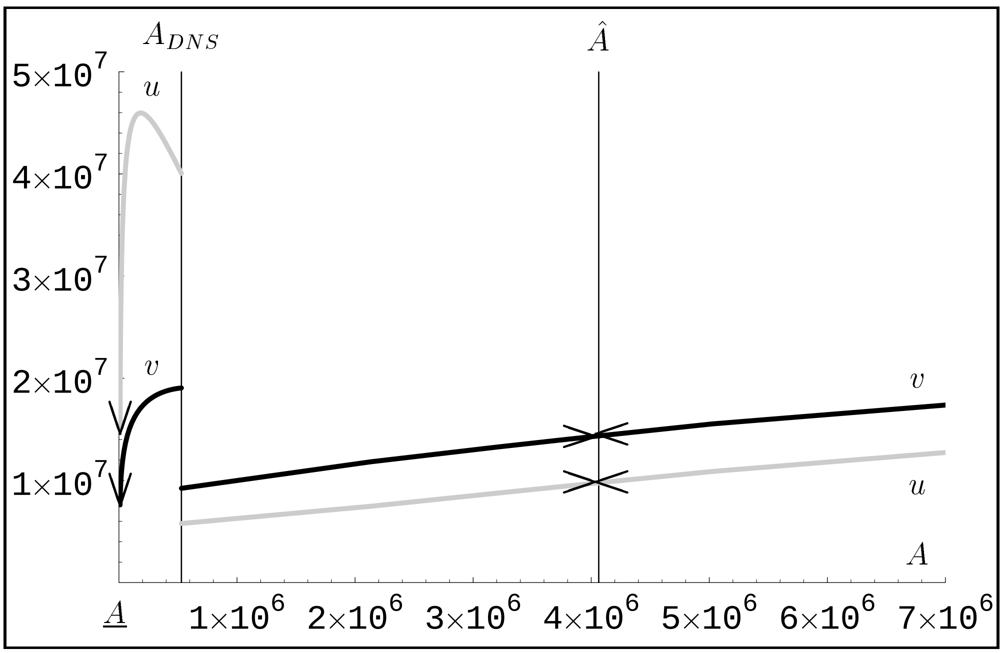
While Tragler et al. (2001) derive a series of interesting policy conclusions, here we focus only on one particular result, i.e. the occurence of a *DNS* point A_{DNS} (follow Fig. 2).[3] This threshold separates the basins of attraction of a low ($\underline{A}$; close to zero) and a high steady state ($\hat{A}$), respectively.[4] The interpretation of

[2] Such diminishing returns reflect 'cream skimming' (cf. Rydell and Everingham, 1994). The image is that some users are easier to treat than others are, and that the treatment system has some capacity and incentive to focus efforts on individuals who are more likely to benefit from an intervention. As treatment funding and the number of people treated grows, the system can afford to be less and less selective.

[3] The numerical value for A_{DNS} calculated by Tragler (1998) is 529,117 users. As far as we know, this is the first reference of a *DNS* threshold in a model validated with empirical data.

[4] The low steady state at $\underline{A}$ is a boundary equilibrium resulting from the pure state constraint $A \geq \underline{A}$. Feichtinger et al. (1997) deal with a similar optimal control problem, for which they provide a complete proof of the optimality of such a boundary steady state and the trajectory leading to it. In their model, this trajectory may in addition hit a second, mixed (i.e. state-control) constraint.

Figure 2: Treatment (grey) and enforcement (black) (both in thousands of dollars) as functions of A along the optimal paths; the left and right vertical lines indicate the *DNS* threshold A_{DNS} and the high steady state value $\hat{A}$, respectively.



the optimal policy is as follows. Ideally society would catch a new drug epidemic early, intervene very aggressively on both the demand (treatment) and supply side (enforcement), and short-circuit the drug outbreak. Very quickly, however, an epidemic can become sufficiently established that the best policy is to moderate the epidemic, rather than to seek eradication.

3 A Model of Drug Enforcement and Property Crime

Dworak (1999) and Caulkins et al. (2000) analyse a model that can be seen as variant to the model presented in Section 2. On the one hand, their model is easier, because it considers only one control, i.e. price-raising law enforcement. On the other hand, Caulkins et al.'s (2000) model is more complex than that by Tragler et al. (2001) in that it takes into account drug enforcement's effects not only on the dynamics of drug use but also on property crime. The motivation is that price-raising enforcement has advantages due to its ability to reduce the number of users and the per capita amounts of consumption (cf. Section 2), while at the same time it has disadvantages, because it may promote property crime. The natural goal hence is to identify conditions under which enforcement's advantages outweigh its disadvantages and vice versa, where it is important to note that Dworak (1999) assumes that drug-involved, convicted property offenders participate in a so-called drug court program.[5]

Let A and v again denote the number of users and law enforcement spending, respectively. Law enforcement and its effect on price is modeled exactly as in Tragler et al. (2001); see Equation (1). With per capita demand being modeled again as $\theta p^{-\omega}$, spending on the drug per user is proportional to $p^{1-\omega}$. It makes sense to assume that the propensity for users to commit a property crime, φ, is a – monotonously increasing – function of the money needed for consumption of the illicit drug, i.e. $\varphi = \varphi\left(p^{1-\omega}\right)$.

In the one model variant by Dworak (1999) which is of interest for this paper, initiation is as given in Equation (2). While the background rate of desistance is as in Equation (3), now there is an additional outflow of the stock of users, i.e. the successfully treated arrestees,

$$\alpha_1 \alpha_2 \varphi\left(p^{1-\omega}\right) A,$$

where α_1 stands for the probability of getting arrested given the commission of a property crime, and α_2 denotes the probability that treatment is successful and removes the criminal from the status of being a drug user.

[5]Drug courts channel drug law offenders into court-supervised treatment programs instead of prisons or jails (Office of National Drug Control Policy, 1998).

Consequently, the evolution of the number of users is described by

$$\dot{A}(t) = kp\left(A(t), v(t)\right)^{-a} A(t)^{\alpha} - \mu p\left(A(t), v(t)\right)^{b} A(t) - \alpha_1 \alpha_2 \varphi\left(p\left(A(t), v(t)\right)^{1-\omega}\right) A(t).$$

The objective functional consists of four terms, which are the social costs of drug use, the social costs of drug-related property crime, the costs of the drug court program, and the enforcement costs. In mathematical terms, the objective is specified as

$$\min_{v(t)\geq 0} \int_0^{\infty} e^{-rt} \left(\underbrace{\kappa\theta A(t) p\left(A(t), v(t)\right)^{-\omega}}_{\text{drug use}} + \underbrace{\beta\varphi\left(p\left(A(t), v(t)\right)^{1-\omega}\right) A(t)}_{\text{property crime}} + \right.$$
$$\left. + \underbrace{\alpha_1 d\varphi\left(p\left(A(t), v(t)\right)^{1-\omega}\right) A(t)}_{\text{drug court program}} + \underbrace{v(t)}_{\text{enforcement}} \right) dt,$$

where r is the discount rate, κ is the social cost per unit of consumption, β is the social cost per property crime, and d represents the costs of an arrest and an ensuing drug court program.

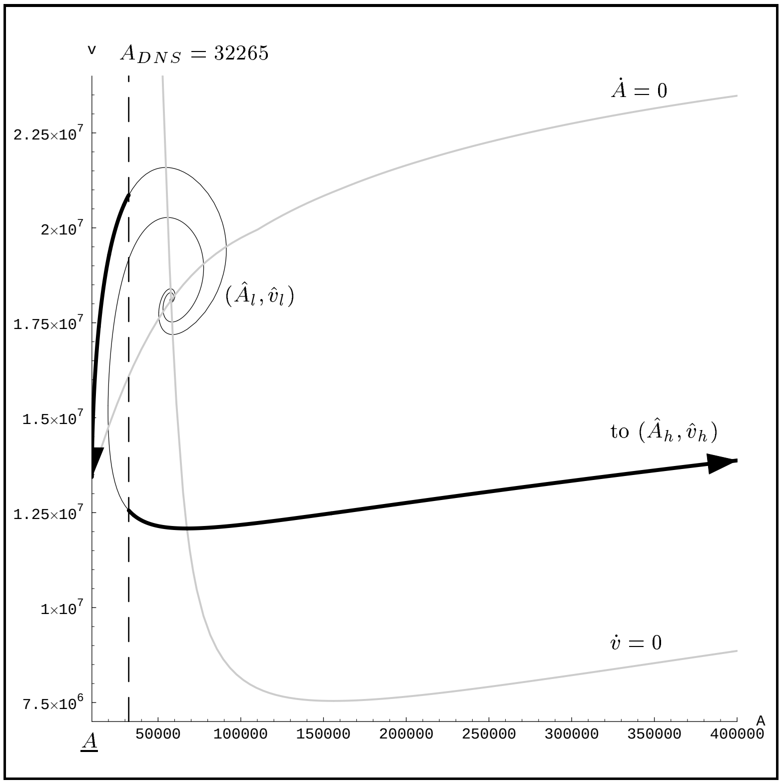
With parameters derived again from data of the current U.S. cocaine epidemic, the solution of this model is as illustrated in Fig. 3, i.e. it exhibits a *DNS* threshold akin to the model presented in Section 2. One might want to argue that – due to the similarities of the two models discussed so far – the history-dependence of the optimal solution of the property-crime model is not a surprising result. The mathematical structure of these two models, however, is sufficiently different so that from this point of view the occurrence of a Skiba point is not obvious a priori. Further discussion is deferred to the concluding Section 5.

4 Optimal Control of Methadone Treatment in Preventing Blood-Borne Disease

While the two models described in the preceding sections share a significant number of properties, the model we deal with in this section is of completely different structure. The purpose is to investigate the influence of methadone maintenance treatment (MMT) on the spread of blood-borne diseases like Hepatitis C Virus (HCV) and Human Immunodeficiency Virus (HIV) among injection drug users (IDUs). In particular, the aim is to find the optimal policy to minimize the discounted stream of the overall costs arising from MMT and the social costs caused by new infections from HIV and HCV.

The model is based on Pollack (2000) and divides the drug users into two groups: the shooting gallery participants who face a high infection risk from

Figure 3: For initial values left to A_{DNS}, it is optimal to approach the minimum steady state at $\underline{A}$, while for initial values right to A_{DNS}, the stable manifolds of the saddle at $\hat{A}_h$ yield the optimal path.



needle sharing, and those who use clean needles and face no disease risk. MMT is targeted to both groups equally, which means that it is unknown whether a person under treatment participates in shooting galleries or not.

The full model, which is described in more detail in Balta (2002), has two states, i.e. the number of IDUs N and the number of infected IDUs I. The single control M represents the number of MMT slots (i.e. number of users getting MMT). Let δ, μ, and γ represent the exit rate from the active IDU population, the exit rate from treatment, and the permanent cure rate of treatment, respectively. Then the state equation for N is given by

$$\dot{N}(t) = \theta - (N(t) - M(t))\,\delta - M(t)\mu\gamma,$$

where θ is the exogenous inflow of drug users, $(N - M)\,\delta$ is the outflow of users not in MMT, and $M\mu\gamma$ is the outflow of users getting MMT. Clearly, $\mu\gamma > \delta$, i.e. the desistance rate is higher for users in treatment.

With κ, λ, and Ω denoting the infectivity, the arrival rate into shooting galleries, and the proportion of shooting gallery participants, respectively, the number of infected IDUs I evolves over time as follows:

$$\dot{I}(t) = \kappa\lambda\left(1 - \frac{M(t)}{N(t)}\right)(\Omega N(t) - I(t))\,\frac{I(t)}{\Omega N(t)} - \delta\left(1 - \frac{M(t)}{N(t)}\right)I(t) - \mu\gamma\frac{M(t)}{N(t)}I(t),$$

where $\left(1 - \frac{M}{N}\right)(\Omega N - I)$ are the candidate users to be infected, $\frac{I}{\Omega N}$ is the probability of sharing a needle with an infected user, $\delta\left(1 - \frac{M}{N}\right)I$ is the outflow of infected users getting no MMT, and $\mu\gamma\frac{M}{N}I$ is the outflow of infected users receiving MMT.

The objective in this model is to choose M so as to minimize the discounted stream of the sum of the overall costs of MMT, the social costs arising from new HIV/HCV infections ('incidence model'),[6] and the social costs of injection drug use (in particular, heroin use). This results in

$$\min_{0 \le M(t) \le N(t)} \int_0^\infty e^{-rt}\left[g\left(M(t)\right) + \nu\kappa\lambda\left(1 - \frac{M(t)}{N(t)}\right)(\Omega N(t) - I(t))\,\frac{I(t)}{\Omega N(t)} + \rho N(t)\right]dt,$$

where r is the discount rate, $g(M)$ denotes the amount of money spent on MMT,[7] ν is the lifetime social cost per new infected user, and ρ describes the daily social cost per drug user.

[6]A parallel study by Gavrila et al. (2002) assumes current costs ('prevalence model').

[7]It makes economic sense to assume that $g(.)$ is a convex function for which it holds that $g(0) = 0$, $g' > 0$, $g'' > 0$. This reflects that on every additional user more money on MMT needs to be spent in order to achieve the same result. The intuition is that it is more difficult to find a user who is motivated to get into MMT if the group of people getting this treatment is already large. Note that this is an alternative way to model the 'cream skimming' effect as described in Section 2 (cf. Footnote 1).

Whereas the analysis of the full model including both states N and I has not yet been carried out completely, Balta (2002) recently completed the analysis for a simplified version of the model, in which it is assumed that the number of users is constant for all times, i.e.

$$N(t) = \bar{N} = \text{const.} \quad \forall t. \tag{4}$$

On the one hand, this assumption does not look too restrictive, because a constant number of users is the best description for a drug problem which is in steady state. On the other hand, the assumption of a constant N does not allow to investigate the full benefits of MMT with respect to its original use, which is the reduction of drug use. In other words, by making the assumption (4), we restrict the effects of MMT mainly to the reduction of the spread of blood-borne diseases, where methadone operates like a perfectly effective needle exchange or heroin maintenance program. The reduction of drug use due to MMT is taken into account only for the infected IDUs, but not for those IDUs who are not infected.

For the basic parameter values, the optimal solution is as illustrated in Fig. 4.[8] Here, $J = \frac{I}{N}$ is the fraction of infected persons in the IDU population, and $U = \frac{M}{N}$ denotes the fraction of IDUs receiving MMT, so the feasible region reduces to $[0,1] \times [0,1]$.[9] We see that the optimal path leads to the saddle point equilibrium x_S, where almost no IDUs are infected.

One may ask the question why – for all initial states – it is optimal to almost 'eradicate' the infections among the IDU population by providing high levels of MMT slots. The answer is that this is due to the high infection risk and the high costs of HIV and HCV treatment in comparison with the costs of MMT. However, we also see that besides the saddle-point steady state x_S and the uncontrolled equilibrium at J_{unc}, in addition there exists another – interior – steady state x_F, which is an unstable focus. As already stated in the introduction and illustrated in the description of the first two models, unstable foci frequently give rise to the occurrence of *DNS* thresholds. So we may expect that a Skiba point arises also in this model, if the parameter values are changed appropriately.

An ideal candidate for causing such a qualitative change in the optimal solution is the parameter ν, which gives the lifetime social cost per new infection. In particular, if a Skiba point is to arise, we expect it to separate the basins of attraction of the low steady state x_S and the uncontrolled one at J_{unc}. We know that in the base case (Fig. 4) the high costs of HIV and HCV treatment, which are themselves a major part in the parameter ν, justify 'eradication' of infections among IDUs for all initial states. This suggests that if ν becomes sufficiently small and hence makes the costs of MMT comparably expensive, it may become

[8]Fig. 4 is for the case of HIV; the optimal solution for the case of HCV looks qualitatively the same.

[9]Indeed, the model only makes sense for values of J below the uncontrolled steady state level J_{unc}, which is the point of intersection of the $\dot{J} = 0$ isocline and the abscissa $U = 0$.

Figure 4: Optimal solution of the MMT model for base case parameter values; for all initial values, it is optimal to converge to the saddle-point steady state close to zero.

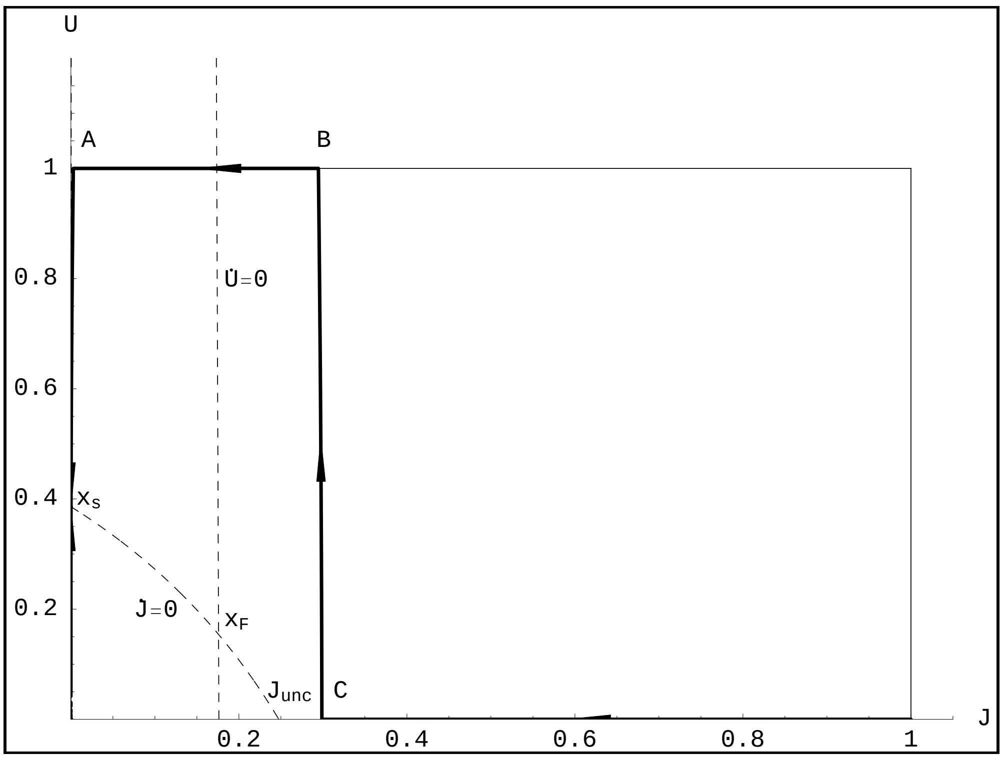
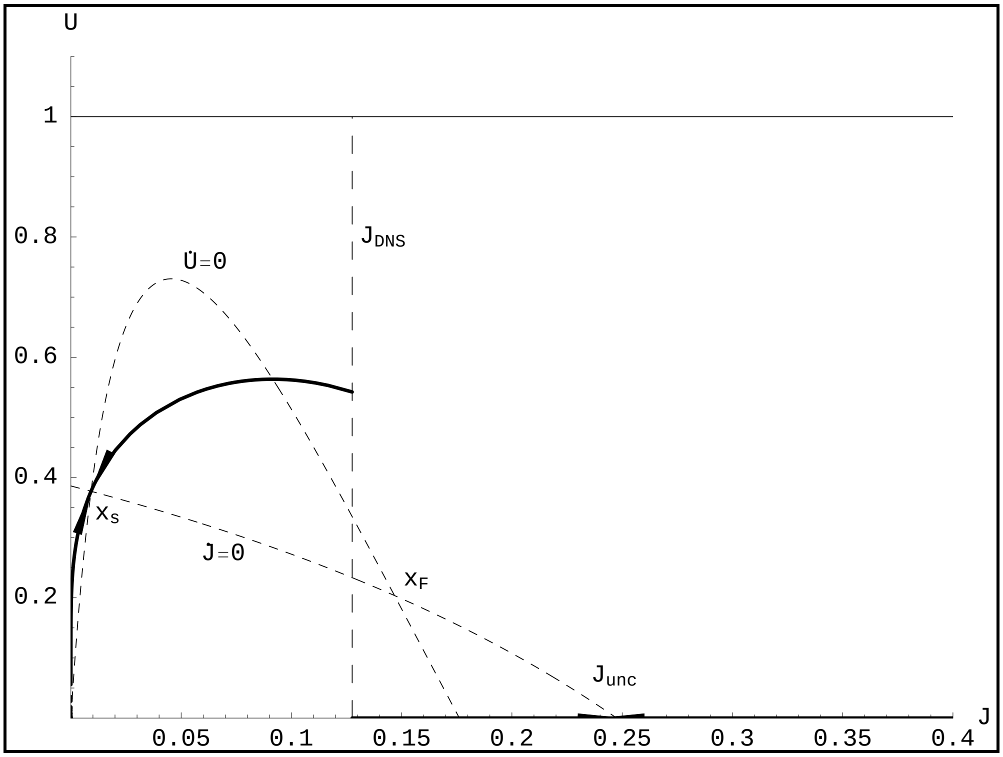


Figure 5: Optimal solution of the MMT model for a value of ν lower than in the base case scenario; the *DNS* threshold J_{DNS} separates the basins of attraction of the low saddle-point steady state x_S and the uncontrolled equilibrium at J_{unc}.



optimal not to provide MMT, at least for high enough fractions of infected IDUs. Fig. 5 demonstrates that this is indeed the case.

5 Concluding Remarks

Recent work in dynamic economics has stressed the fact that economic outcome may be history-dependent. Considering a convex-concave production function in a Ramsey-type optimal growth model, Skiba (1978) stressed the fact of multiple steady states. In that case, two saddle-points and one unstable steady state in-between occur. Multiple equilibria constitute a low-level, manageable form of complexity, i.e. the non-uniqueness of the optimal control in the Skiba threshold. A first sound proof of the existence of a threshold separating two basins of attraction in a one-state dynamic optimization model was given by Dechert and Nishimura (1983). Since that time, many papers have been written dealing with multiple steady states and history-dependence in one-dimensional optimal control models; see, e.g., Feichtinger and Wirl (2000), and Deissenberg et al. (2001) for some interesting results. The thresholds separating the equilibria have been called 'Skiba points' by Brock and Malliaris (1989, Chap. 6); see also Feichtinger and Hartl (1986, pp. 116-119, 325-335).

In this paper, we focused on three particular one-state optimal control models in the context of illicit drug consumption, all of which were validated with empirical data and exhibit Skiba points. A key question is, if these three models share certain properties, which give rise to the occurrence of such thresholds. Indeed, we find one and the same mechanism in each of the three models, which seems to be particularly responsible for this property. This mechanism is the state dependency of the inflow term into the state, which is $kp^{-a}A^{\alpha}$ in the first two models and $\kappa\lambda\left(1-\frac{M}{N}\right)(\Omega N-I)\frac{I}{\Omega N}$ in the third one, respectively. More precisely, there is a *positive-feedback* effect, i.e. the inflow terms are – at least initially – monotonously increasing in the state with the additional property that these terms are zero if the state is zero.

To better understand how this mechanism works, it is further important to note that in all three models the state can be perceived as being 'bad' in the sense that it increases the costs in the objective functional, so small values of the states are desirable. However, due to the endogenous forces that cause the state to increase, keeping the state low or even decreasing it is difficult and hence costly. Consequently, if at all it turns out to be optimal to choose this path of 'eradication', this will only be the case when the state is sufficiently low, i.e. the positive-feedback effect is still small. The larger the state becomes, the more control is needed and hence the more costly it becomes to compensate for the growing feedback effect. Once the state is too big, eradication is no longer justifiable because the increasing control costs overcome the benefits one would have from converging to a low steady state; this is when the Skiba threshold occurs, implying that for values above it it is optimal to follow the path leading to the high steady state.

These somewhat heuristic arguments are supported by an analysis carried out by Tragler (1998) who showed that the Skiba threshold in the first model disappears if the initiation term (2) is replaced by the term $\tilde{k}p^{-a}$, i.e. the feedback effect is omitted. In that case, the inflow term is constantly high even if the state is small, so in the optimal solution there is always convergence to a comparably high steady state. Summarizing, the key mechanism that seems to lead to a Skiba point is the monotonously increasing positive-feedback effect in the inflow terms.[10]

In addition, Tragler (1998) identified a necessary (not sufficient) condition for the occurence of a Skiba point in the model described in Section 2, which is that

[10]In the first two models, this effect is enhanced by the assumption that the control becomes less efficient if the state increases ('enforcement swamping'; see Equation (1) and the description below it).

the controls are allowed to become high enough to make eradication possible. In particular, Tragler (1998) investigated the model also with the additional budget constraint $u + v = GA$, which implies that the available drug control budget is proportional to the size of the drug problem, assuming that the number of users A is a reasonable measure for the size of the problem.[11] If this constraint is imposed, the budget is too small for low numbers of users for eradication being still possible, so the Skiba point disappears.

This raises another interesting, more policy-oriented issue. We know from political practice that budget constraints like the one above are probably more realistic than the assumption that policy makers have unrestricted resources. Consequently, looking at the budget on the optimal paths to the low steady state (in particular, Figs. 2 and 3) it becomes clear that eradication carries a heavy political burden. For it to succeed, policy makers must have the political capital needed to obtain massive funding – even if the specific drug problem is still relatively small – and the will to do so, although they may never receive recognition for averting a drug epidemic that remained invisible to the public eye. Apart from that, it is questionable whether a policy that many people would identify as 'law and order' policy[12] is desirable in a modern society.

Another interesting and probably even macabre observation follows from what has already been discussed partly in Section 4 for the MMT model. We saw that there a history-dependent policy may occur, if the social costs of a new infection become sufficiently low. A reduction in these costs may be obtained for example by achieving a reduction in the costs of HIV/HCV treatment, which is in principle a desirable goal. However, the consequence is that then MMT becomes comparably costly, and the occurrence of the Skiba point implies that for proportions of infected users above the threshold it is no longer optimal to provide MMT, which would help the drug users. Of course, this somewhat perverse effect follows from the purely economic analysis, which neglects aspects like pain and suffering from drug users and their families, friends, etc.

We conclude this paper with the following final remark. Clearly, the existence of multiple equilibria is not restricted to one-state optimal control models. For higher-dimensional state spaces, there may exist Skiba curves or *DNS* surfaces separating various basins of attraction. Until now, however, the existence of such threshold sets can be established only numerically. Caulkins et al. (2001) and Haunschmied et al. (2000,2001) delivered interesting examples of two-dimensional control models. The full MMT model dealt with in Section 4 provides an interesting example in the control of illicit drugs. A preliminary analysis of this two-state MMT optimal control model – $N(t)$ and $I(t)$ being the state variables, and $M(t)$ the control – reveals that there are four stationary points. In particular, it can be

[11] In 1992, the per-capita budget G in the U.S. was around \$ 1,600.

[12] In the first two models, the law enforcement spending required for successful eradication is occasionally more than 200% of actual spending in the U.S. for numbers of users that are far below 10% of the actual number of users.

shown that there exist two interior saddle points, one boundary and one unstable equilibrium. We presume that there exists a *DNS* curve separating the basins of attraction, which to find is part of actual research.

References

Andvig, J.C. (1991) "The economics of corruption: a survey," *Studi Economici* **43**, 57-94.

Balta, J. (2002) *Optimal Control of Methadone Treatment in Preventing Blood-Borne Disease*, Master Thesis, Institute for Econometrics, Operations Research and Systems Theory, Vienna University of Technology.

Becker, G.S. (1968) "Crime and punishment: an economic approach," *Journal of Political Economy* **76**, 169-217.

Brock, W.A. and A.G. Malliaris (1989) *Differential Equations, Stability and Chaos in Dynamical Economics* (North Holland, Amsterdam).

Caulkins, J.P. (1993) "Local drug markets' response to focused police enforcement," *Operations Research* **41**(5), 848-863.

Caulkins, J.P., M. Dworak, G. Feichtinger and G. Tragler (2000), "Price-raising drug enforcement and property crime: a dynamic model," *Journal of Economics* **71**(3), 227-253.

Caulkins, J.P., R.F. Hartl, G. Tragler and G. Feichtinger (2001), "Why politics makes strange bedfellows: dynamic model with DNS curves," *Journal of Optimization Theory and Applications* **111**(2), 237-254.

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

Dechert, W.D. and K. 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.

Deissenberg, C., G. Feichtinger, W. Semmler and F. Wirl (2001) "History dependence, multiple equilibria, and global dynamics in efficiently governed economies," Research Report 258, Institute for Econometrics, Operations Research and Systems Theory, Vienna University of Technology; forthcoming in W. Barnett, C. Deissenberg and G. Feichtinger (eds.), *Economic Complexity: Non-Linear Dynamics, Multi-Agents Economies, and Learning* (Cambridge University Press, Cambridge).

Dworak, M. (1999) *A Dynamic Model of Drug Enforcement and Property Crime*, Master Thesis, Institute for Econometrics, Operations Research and Systems

Theory, Vienna University of Technology.

Feichtinger, G., W. Grienauer and G. Tragler (1997) "Optimal dynamic law enforcement," Research Report 197, Institute for Econometrics, Operations Research and Systems Theory, Vienna University of Technology; forthcoming in *European Journal of Operational Research*.

Feichtinger, G. and R.F. Hartl (1986) *Optimale Kontrolle ökonomischer Prozesse – Anwendungen des Maximumprinzips in den Wirtschaftswissenschaften* (Walter de Gruyter, Berlin).

Feichtinger, G. and F. Wirl (2000) "Instabilities in concave, dynamic, economic optimization," *Journal of Optimization Theory and Applications* **107**, 277-288.

Gavrila, C., H.A. Pollack, J.P. Caulkins, P.M. Kort, G. Feichtinger and G. Tragler (2002) "Optimal control of harm reduction in preventing HIV and Hepatitis C among drug users;" in preparation.

Haunschmied, J.L., G. Feichtinger, R.F. Hartl and P.M. Kort (2001) "Keeping up with the technology pace: A DNS-curve in a limit cycle in a technology investment decision problem," Research Report 262, Institute for Econometrics, Operations Research and Systems Theory, Vienna University of Technology.

Haunschmied, J.L., P.M. Kort, R.F. Hartl and G. Feichtinger (2000) "A DNS-Curve in a two state capital accumulation model: a numerical analysis. Research Report 242, Institute for Econometrics, Operations Research and Systems Theory, Vienna University of Technology; forthcoming in *Journal of Economic Dynamics and Control*.

Kleiman, M.A.R. (1993) "Enforcement swamping: a positive-feedback mechanism in rates of illicit activity," *Mathematical and Computer Modelling* **17**(2), 65-75.

Office of National Drug Control Policy (1998) *The National Drug Control Strategy: 1998* (The White House, Washington, D.C.).

Pollack, H.A. (2000) "The cost-effectiveness of methadone in preventing blood-borne disease: a comparison of HIV and Hepatitis C," Working Paper.

Reuter, P. (1983) *Disorganized Crime: The Economics of the Visible Hand* (MIT Press, Cambridge, MA).

Reuter, P. and M.A.R. Kleiman (1986) "Risks and prices: an economic analysis of drug enforcement," in M. Tonry and N. Morris (eds.), *Crime and Justice: An Annual Review of Research* **7** (University of Chicago Press, Chicago), 289-340.

Rydell, C.P. and S.S. Everingham (1994) *Controlling Cocaine – Supply versus Demand Programs* (RAND, Santa Monica).

Schelling, T. (1973) "Hockey helmets, concealed weapons, and daylight saving: a study of binary choices with externalities," *Journal of Conflict Resolution*, 381-428.

Schelling, T.C. (1978) *Micromotives and Macrobehavior* (W.W. Norton & Company, New York).

Skiba, A.K. (1978) "Optimal growth with a convex-concave production function," *Econometrica* **46**, 527-539.

Tragler, G. (1998) *Optimal Control of Illicit Drug Consumption: Treatment versus Enforcement*, PhD Thesis, Institute for Econometrics, Operations Research and Systems Theory, Vienna University of Technology.

Tragler, G., J.P. Caulkins and G. Feichtinger (2001) "Optimal dynamic allocation of treatment and enforcement in illicit drug control," *Operations Research* **49**(3), 352-362.