

Dynamic Drug Policy: Optimally Varying the Mix of Treatment, Price-Raising Enforcement, and Primary Prevention Over Time

Jonathan P. Caulkins, Gernot Tragler

Research Report 2016-01

February, 2016

Operations Research and Control Systems

Institute of Statistics and Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Wiedner Hauptstraße 8 / E105-4
1040 Vienna, Austria
E-mail: orcocos@tuwien.ac.at

Dynamic Drug Policy: Optimally Varying the Mix of Treatment, Price-Raising Enforcement, and Primary Prevention Over Time

Jonathan P. Caulkins¹, Gernot Tragler^{2,*}

February 6, 2016

¹ H. John Heinz III College, Carnegie Mellon University

² Institute of Statistics and Mathematical Methods in Economics, TU Wien

Abstract

A central question in drug policy is how control efforts should be divided among enforcement, treatment, and prevention. Of particular interest is how the mix should vary dynamically over the course of an epidemic. Recent work considered how various pairs of these interventions interact. This paper considers all three simultaneously in a dynamic optimal control framework, yielding some surprising results. Depending on epidemic parameters, one of three situations pertains. It may be optimal to eradicate the epidemic, to "accommodate" it by letting it grow, or to eradicate if control begins before drug use passes a DNSS threshold but accommodate if control begins later. Relatively modest changes in parameters such as the perceived social cost per unit of drug use can push the model from one regime to another, perhaps explaining why opinions concerning proper policy diverge so sharply. If eradication is pursued, then treatment and enforcement should be funded very aggressively to reduce use as quickly as possible. If accommodation is pursued then spending on all three controls should increase roughly linearly but less than proportionally with the size of the epidemic. With the current parameterization, optimal spending on prevention varies the least among the three types of control interventions.

*Corresponding author. Address: Wiedner Hauptstr. 8/105-4, A-1040 Wien, Austria. Email: gernot.tragler@tuwien.ac.at. Fax: +43-1-58801-910542. Phone: +43-1-58801-10542.

Keywords: Dynamic programming/optimal control: applications to drug policy; Judicial/legal: crime and drug policy.

1 Introduction

Illicit drugs impose enormous costs on society (Harwood et al. [31], UN-ODC [58]), and there is considerable debate over how policy makers should respond. A central question concerns the relative roles of three broad strategies: enforcement, treatment, and prevention.

Drug use varies dramatically over time in ways that can fairly be described as epidemics even though there is no literal pathogen (Golub and Johnson [26], Ferrence [24], Caulkins [5],[7]). For example, cocaine initiation in the US increased roughly four-fold in the 1970s, then the "infectivity" (number of new initiates recruited per current user) subsequently fell over time (Caulkins et al. [8]).

Traditionally drug control effectiveness has been evaluated in a static framework (e.g., Rydell and Everingham [51]), but intuitively the relative roles of enforcement, treatment, and prevention should vary over the course of an epidemic. Indeed, this has been argued for various pairs of interventions (Behrens et al. [2], Caulkins et al. [9], Tragler et al. [57]). The present paper yields substantial new insights by simultaneously considering key elements of all three principal classes of drug control interventions in a dynamic model parameterized for the most problematic drug (cocaine) for the country with the most dependent users (the US).

Enforcement, treatment, and prevention are broad classes of interventions, not single programs, so it is important to clarify what specifically is modeled. Enforcement here refers to actions taken against the drug supply chain that raise the cost of producing and distributing drugs and thereby increase retail prices (cf., Reuter and Kleiman [46]). Such actions account for the majority of US enforcement spending. For enforcement within US borders the largest cost driver is incarceration. Simply put, prison (at \$25-30,000 per cell-year) costs more than arrest or adjudication (Greenwood et al. [28]). More people are arrested for possession than sale, but on the order of 90+% of those imprisoned for drug-law violations in the US were involved in drug distribution (Sevigny and Caulkins [54]).³

³Possession arrests include "possession with intent to distribute", which is essentially a distribution charge, but offenders arrested for simple possession are less likely to be incarcerated and when they are, they serve shorter sentences. Note that many of those involved in distribution also use drugs, but generally it is not the use per se that leads to their incarceration.

A smaller share of enforcement dollars are spent outside US borders on interdiction in source countries and the "transit zone". There is debate concerning whether these activities are best thought of as driving up equilibrium prices or as creating spot shortages (Rydell and Everingham [51], Crane et al. [17], Manski et al. [39], Caulkins et al. [9]). Modeling price raising enforcement is of interest even if enforcement outside the US has no impact on equilibrium prices, but we suspect that it does have at least some such effects.

Enforcement has been hypothesized to work through other mechanisms as well. Moore [42] and Kleiman [36] suggest it might increase non-monetary "search costs" that users incur to obtain drugs. These costs are non-negligible, even for experienced users (Rocheleau and Boyum [49]), but since regular users often have 10-20 alternative suppliers (Riley [48]) enforcement's effects through increasing search time are second-order for established markets (Caulkins [3]) such as those for cocaine in the US today.⁴ Likewise, enforcement against suppliers of mass market drugs does not work primarily through incapacitation; there are few barriers to entry, so incarcerated sellers are rapidly replaced (Kleiman [37]).

Prevention is similarly multi-faceted. Unfortunately there is little scientific evidence concerning the effectiveness of most forms of prevention other than school-based prevention (Cuijpers [18]), so we focus on school-based programs and adapt parameter estimates from Caulkins et al. [13],[14].

Caulkins et al.'s estimates are based on lifetime projections of results for "best practice" programs evaluated in randomized control trials run through the end of high school. This has two implications. First, since data are only available on impacts through the end of high school, there is unavoidable uncertainty about prevention's effectiveness over a lifetime. Second, the estimates pertain to model programs. Historically most school districts have not implemented research-based programs with high fidelity (Hallfors and Godette [30]). By using Caulkins et al.'s data, we are examining what the optimal level of spending on school-based prevention would be if the best currently available prevention technologies were employed.

There are many kinds of treatment, and they are of varying quality (Institute of Medicine [33],[34]). Effectiveness data from randomized-controlled trials for cocaine treatment is lacking (Manski et al. [40]). Hence, we model treatment somewhat abstractly as simply increasing the net quit rate and ignore the possibility that it might reduce the social damage per unit of

⁴Infrequent or "light" users may have fewer alternative suppliers, but they account for a modest share of all consumption because they use so much less, per capita, than do heavier users.

consumption. For consistency we use the same basecase assumptions about treatment’s average cost and effectiveness as did Rydell and Everingham [51] and Tragler et al. [57], but in light of Manski et al. we do sensitivity analysis with respect to those assumptions.

Note that our goal is not to anoint any one of these classes of interventions as the ”winner” in some cost-effectiveness horse race. Rather, the goal is to understand better how their relative roles might vary over the course of an epidemic.

2 The Model

2.1 Clarifying Some Common Misconceptions

Before proceeding it is important to dispel some common misconceptions about drug markets. First, most new users are introduced to drugs by other users, typically friends or siblings. This is the sense in which drug use is ”contagious”. Dealers rarely ”push” drugs on unwitting innocents (Kaplan [35]). Furthermore, drug supply is characterized by intense and atomistic competition (Reuter [45]), not monopolistic control. Hence, drug suppliers do not act strategically. There are simply too many of them; well over a million Americans sold cocaine within only twelve months (Caulkins [6]).⁵ Hence, one can develop sensible models of drug markets without explicitly modeling strategic behavior by suppliers. Instead, one can simply abstract the drug supply sector by what amounts to a supply curve (albeit one whose position depends on enforcement).

Second, drug initiation and use are affected by prices. There was once a lore that drug addicts ”had to have their drug” regardless of the price, but a considerable literature has clearly established that cocaine use responds to price changes (Grossman and Chaloupka [29], Chaloupka et al. [15], Chaloupka and Pacula [16], Rhodes et al. [47], DeSimone [21], DeSimone and Farrelly [22], Dave [19]). Gallet [25] provides a nice, new literature review and synthesis. This should not be surprising. Merely consuming less when prices rise in no way implies or requires perfect foresight or full rationality. What is somewhat surprising is the magnitude of the response. Best estimates for the elasticity of demand for cocaine are in the neighborhood

⁵Market power is most concentrated at the export level in Colombia, and never more so than in the heyday of the Medellin ”cartel”. Yet this supposed ”cartel” was not able to stave off a precipitous decline in prices. In reality, the cartel was formed more for protection against kidnapping than to strategically manipulate prices. Today there are several hundred operators even at that market level.

of -1 (Caulkins and Reuter [10]), implying that a one percent increase in price is associated with a one percent reduction in use. This substantial responsiveness may stem from the fact that the vast majority of cocaine is consumed by dependent users who spend a large share of their disposable income on their drug of choice. All other things being equal, price elasticities tend to be larger for things that are important budget items (e.g., housing) than for incidentals (e.g., toothpaste).

Unfortunately, there is much less information concerning what proportion of the overall elasticity stems from reduced per capita consumption by existing users vs. reduced initiation or increased quitting changing the number of users. In the absence of better information, we follow Rydell and Everingham [51] and Tragler et al. [57] in assuming an equal division between these categories and likewise divide the latter (price elasticity of prevalence) equally between effects on initiation and quitting.

2.2 Model Structure

The present model extends that of Tragler et al. [57]. It tracks the number of users ($A(t)$) over time t . Initiation is modeled as an increasing (but concave) function of the current number of users that is modified by price, through a constant price elasticity of initiation, and by prevention.

Primary prevention is typically modeled as reducing initiation by a certain percentage, where the percentage depends on program intensity. Diminishing returns are presumed through an exponential decay as in Behrens et al. [2]. As mentioned, effectiveness estimates are based on Caulkins et al.'s [13],[14] analysis of "model" or "best practice" programs. Note: even "model" prevention is no panacea. As Caulkins et al. observe, prevention tends to be cost-effective primarily because it is so cheap, not because it is extremely effective. If kids who were going to initiate drug use in the absence of a prevention intervention are given cutting edge school-based drug prevention, most (though not all) would still initiate drug use. That does not necessarily mean prevention programs are poorly designed. It may simply indicate that there is little one can possibly do in 30 or so school contact-hours to counteract the influence of many thousands of hours of television, peers, etc.

The background quitting rate is assumed to be a simple constant per capita rate. (Even such simple modeling can fit historical data surprisingly well; cf., Caulkins et al. [8].) Like initiation, this flow is affected by price through a constant elasticity and by an intervention, in this case treatment. As in Rydell and Everingham [51] and Tragler et al. [57], treatment is assumed to exhibit diminishing returns because some users are more likely to relapse than others, and the treatment system has some capacity to target

interventions first on those for whom the prognosis is most favorable.

Price is a function of enforcement intensity. The underlying theoretical paradigm is Reuter and Kleiman’s [46] ”risks and prices” model, operationalized as in Caulkins et al. [12]. The key insight is that some component of price (the intercept) is due to the ”structural consequences of product illegality” (Reuter [45], Caulkins and Reuter [11]) accompanied by some minimal enforcement. The increment in price above that intercept is driven by the *intensity*, not the *level*, of enforcement because of ”enforcement swamping” (Kleiman [37]). Sellers do not care per se about the level of enforcement, e.g., the number of arrests. They care about their individual arrest risk, which is essentially the total number of arrests divided by the number of sellers subject to those arrests. Hence, for any given level of enforcement, the *intensity* is inversely related to the number of sellers. Since we do not model sellers explicitly, we divide by the number of users, implicitly assuming that the number of sellers is proportional to the number of users.

We assume that the social planner wishes to minimize the discounted weighted sum of drug use and of drug control spending. The cost coefficient on consumption is simply the average social cost per unit of cocaine use. Clearly marginal costs would be more relevant, but we have no way to estimate them.

The quantity of cocaine consumed is simply the number of users times the baseline consumption per user, adjusted for the short-term price elasticity of consumption per capita. Consumption per capita varies across users and the mix of light and heavy users varies over the course of an epidemic. Our consumption per capita is calibrated to our base year (1992), a time when roughly one-third of all users were heavy users (weekly or more often).

2.3 Mathematical Formulation

If we let $u(t)$, $v(t)$, and $w(t)$ denote treatment, enforcement, and prevention spending, respectively, then the discussion above suggests the following formulation:

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

subject to

$$\begin{aligned} \dot{A}(t) = & kA(t)^\alpha p(A(t), v(t))^{-a} \Psi(w(t)) - c\beta(A(t), u(t))A(t) - \\ & - \mu p(A(t), v(t))^b A(t) \end{aligned}$$

and the non-negativity constraints

$$u(t) \geq 0, v(t) \geq 0, w(t) \geq 0,$$

where

J = discounted weighted sum of the costs of drug use and control,
 r = time discount rate,
 κ = social cost per unit of consumption,
 θ = per capita rate of consumption at baseline prices,
 $A(t)$ = number of users at time t ,
 $p(A(t), v(t))$ = retail price,
 ω = absolute value of the short-run price elasticity of demand,
 k = constant governing the rate of initiation,
 α = exponent governing concavity of contagious aspect of initiation,
 a = absolute value of the elasticity of initiation with respect to price,
 $\Psi(w(t))$ = proportion of initiation remaining after prevention,
 c = treatment efficiency proportionality constant,
 $\beta(A(t), u(t))$ = outflow rate due to treatment,
 μ = baseline per capita rate at which users quit without treatment, and
 b = elasticity of desistance with respect to price.

As in Tragler et al. [57], treatment's increment to the per capita outflow rate is assumed to be proportional to treatment spending per capita raised to an exponent (z) that reflects diminishing returns, with a small constant in the denominator (δ) to prevent division by zero:

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

We model enforcement's effect on price as in Caulkins et al. [12] and Tragler et al. [57]:

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

where d describes the price with minimal enforcement, e is the enforcement efficiency proportionality constant, and ϵ is an arbitrarily small constant that avoids division by zero.

Following Behrens et al. [2], we model prevention as reducing initiation by a certain proportion. That proportion increases with prevention spending but at a decreasing rate because of diminishing returns. Specifically, we model

$$\Psi(w(t)) = h + (1 - h)e^{-mw(t)}$$

for positive constants h and m .

2.4 Parameters

Tragler et al. [56] describe in detail how parameters are derived from the literature. Briefly, the price elasticity parameters (a , b , and ω) collectively generate a long term price elasticity of demand of -1 (Caulkins et al. [12], Caulkins and Reuter [10]), half coming from reduced consumption by current users (ω) and half from changes in the number of users, with the latter divided equally between impacts on initiation (a) and quitting (b).

For consistency with Rydell and Everingham [51] and Tragler et al. [57], we take the baseline price to be \$106.73 per pure gram and choose as initiation parameters $\alpha = 0.3$ and $k = 5,167$ to make initiation 1,000,000 per year when the number of users $A = 6,500,000$ in base conditions. They estimate total baseline consumption as 291 (pure) metric tons, so we set $\theta = 14.6259$ (since $14.6259 \times 0.10673^{-0.5} = 291,000,000/6,500,000$ and price is expressed in thousands of dollars).

Rydell and Everingham [51] (p.38) report cocaine-related health and productivity costs of \$19.68B for cocaine in 1992, dividing by 291 metric tons of consumption implies an average social cost per gram of \$67.6/gram (in 1992 dollars). These figures do not include crime-related costs, so in light of Miller et al. [41], we take \$100/gram as our base value ($\kappa = 0.1$ since dollars are measured in thousands). In view of Caulkins et al. [13] we also consider larger values in the sensitivity analysis.

The price function parameters ($d = 0.06792$ and $e = 0.02655$) reflect a price of \$106.73 per gram under base case enforcement spending and an elasticity of price with respect to enforcement spending of 0.3636 as in Caulkins et al. [12].

As in Tragler et al. [57] we assume $c = 0.04323$ and $z = 0.6$. These values reflect Rydell and Everingham's [51] estimates that spending an average of \$1,700 - \$2,000 per admission to treatment provides a 13% chance of ending heavy use, over and above baseline exit rates.

We adopt Behrens et al.'s [2] value of $h = 0.84$, but modify their value of m slightly (1.93×10^{-6} vs. 2.37×10^{-6}) to reflect better the size of the birth cohorts on whom prevention is targetted.

The outflow parameter $\mu = 0.18841$ was selected to make the outflow be 700,000 users per year at base case prices, which reflects the observed population change (ONDCP [43]) net of initiation and treatment during the recent years of relative stability. The discount rate is set at $r = 0.04$ as in Rydell et al. [50] and Caulkins et al. [12].

These values are summarized in Table 1. Two values are given for parameters d , e , k , κ , μ , and θ . The values in brackets are the ones just described. For analytical convenience, we adjust d , e , k , and μ so that $\kappa = 1$ and $\theta = 1$,

Table 1: Base case parameter values

Parameter	Value	Description
a	0.25	absolute value of the elasticity of initiation with respect to price
α	0.3	exponent reflecting contagiousness of initiation
b	0.25	elasticity of desistance with respect to price
c	0.04323	treatment efficiency proportionality constant
d	0.03175 [0.06792]	price with minimal enforcement (in thousands of \$)
δ	0.001	constant to avoid division by zero
e	0.01241 [0.02655]	enforcement efficiency proportionality constant
ϵ	0.001	constant to avoid division by zero
h	0.84	one minus maximum proportion of baseline initiation prevention can avert with full implementation
k	4,272 [5,167]	initiation constant
κ	1 [0.1]	social cost per gram consumed (in thousands of \$)
m	1.93×10^{-6}	prevention efficiency proportionality constant
μ	0.22786 [0.18841]	natural outflow rate from use
ω	0.5	absolute value of the short run elasticity of demand
θ	1 [14.6259]	per capita consumption constant
r	0.04	annual discount rate (time preference rate)
z	0.6	$1 - z$ reflects treatment's diminishing returns

yielding the second set of values for those parameters.

3 Base Case Analysis

Note that for simplicity, the time argument t will mostly be omitted from now on. The model cannot be solved analytically, but the Appendix describes the derivation of the necessary optimality conditions according to Pontryagin's maximum principle (cf. Feichtinger and Hartl [23], Grass et al. [27], Leonard and Long [38]). Due to the concavity of the Hamiltonian with respect to all three controls (u , v , w), setting the first-order partial derivatives equal to zero leads to the unrestricted extremum. These equations allow one to describe u and w as functions of v and A , so the solutions are described in terms of phase portraits in the A - v plane.

Steady state values are given by intersections of the isoclines obtained by setting to zero the derivatives of the state (A) and control (v) variables (dark gray and black curves, respectively, in Figure 1). With parameter values from Table 1, there are two intersections, a left-hand (lower A) intersection ($\hat{A}^{(l)} = 0.2 \times 10^6$, $\hat{v}^{(l)} = 1.04 \times 10^7$) that is an unstable focus and a right-hand (larger A) intersection that is a saddle point ($\hat{A}^{(h)} = 3.24 \times 10^6$, $\hat{v}^{(h)} = 1.14 \times 10^7$). Every saddle point equilibrium in a two-dimensional phase portrait has a stable manifold which consists of two branches. Locally, these branches are determined by the eigenvector associated with the negative eigenvalue of the Jacobian evaluated at the steady state. This is used to numerically compute the complete stable manifolds (light gray curves in Figure 1) which, in optimal control theory, are known to be candidates for the optimal trajectories.

The stable manifold from the right describes directly what trajectory one should follow to drive the number of users down to the saddle point equilibrium if the initial conditions have $A(0) > \hat{A}^{(h)}$. The stable manifold from the left emanates from the unstable focus, so it is not immediately obvious what the optimal policy should be when starting to the left of that focus. If control begins when the number of users is below its steady state value but still above a certain threshold A_{DNSS} to be described shortly, then the optimal treatment, prevention, and enforcement rates gradually increase while $A(t)$ converges to the equilibrium $\hat{A}^{(h)}$. (The opposite holds for initial states above the steady state value, but we presume that control begins with $A(0) < \hat{A}^{(h)}$.) Note this means that even if the optimal policy is pursued, the number of users will increase over time toward the equilibrium ($\hat{A}^{(h)}$).

Figure 1: Phase portrait with base case parameter values. The intersections of the isoclines $\dot{A} = 0$ and $\dot{v} = 0$ give the two steady-state solutions. The light gray curves represent the stable manifolds of the saddle point.

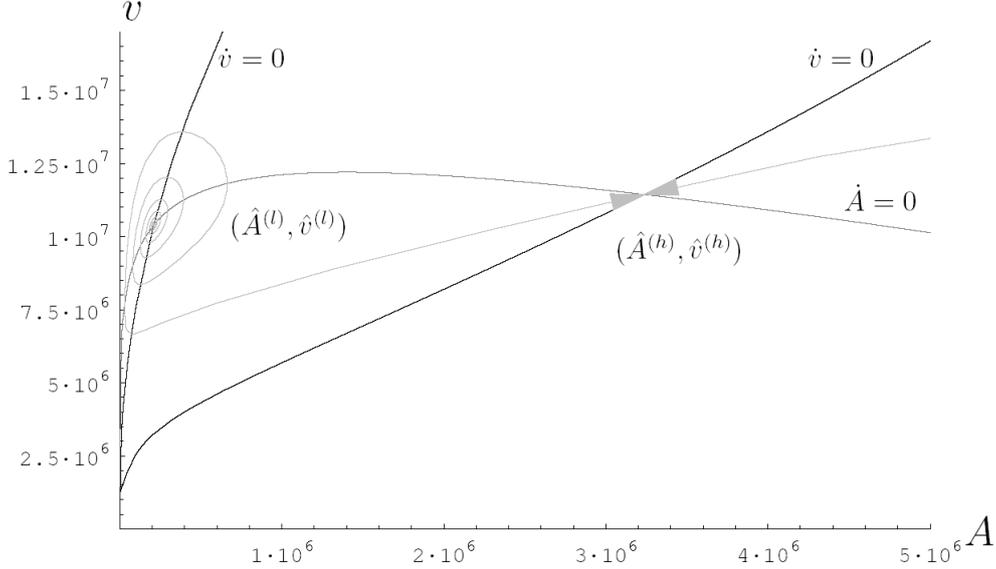
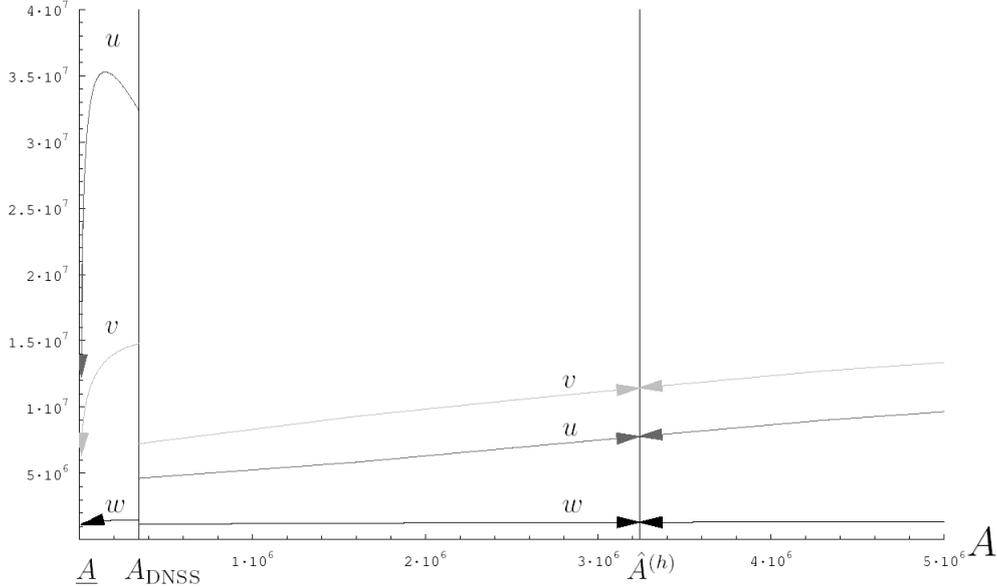


Figure 2 shows the optimal amounts of treatment, prevention, and enforcement spending as functions of the number of users. When $A(0) > A_{DNSS}$, the optimal levels of control spending (u , v , and w) are each approximately linear in the size of the epidemic (A). The treatment (u) and enforcement (v) lines are almost parallel, implying that as time goes by, increments in the treatment and enforcement budgets should be approximately equal. Since with these parameter values the enforcement spending trajectory has a higher "intercept", for $A(0) > A_{DNSS}$ it is always optimal to spend more on enforcement than on treatment, but enforcement's share of the total control budget shrinks as time goes on.

According to Figure 2, spending on prevention should also increase as the epidemic grows but not by much for the simple reason that prevention should already be almost "maximally funded" even when the epidemic is small. "Maximally funded" is in quotes because there is no literal bound on prevention spending, but the least it is ever optimal to spend on prevention is about \$1B per year. A cutting edge junior-high school-based prevention program costs about \$150 per youth, even including "booster sessions" in the two subsequent years (Caulkins et al. [13]), so \$1B per year would be enough to give 6 million youth per year an excellent prevention program. Since there

Figure 2: Treatment (dark gray), enforcement (light gray), and prevention (black) as functions of A along the optimal paths. The left and right vertical lines represent the DNSS threshold and the saddle point at $\hat{A}^{(h)}$, respectively.



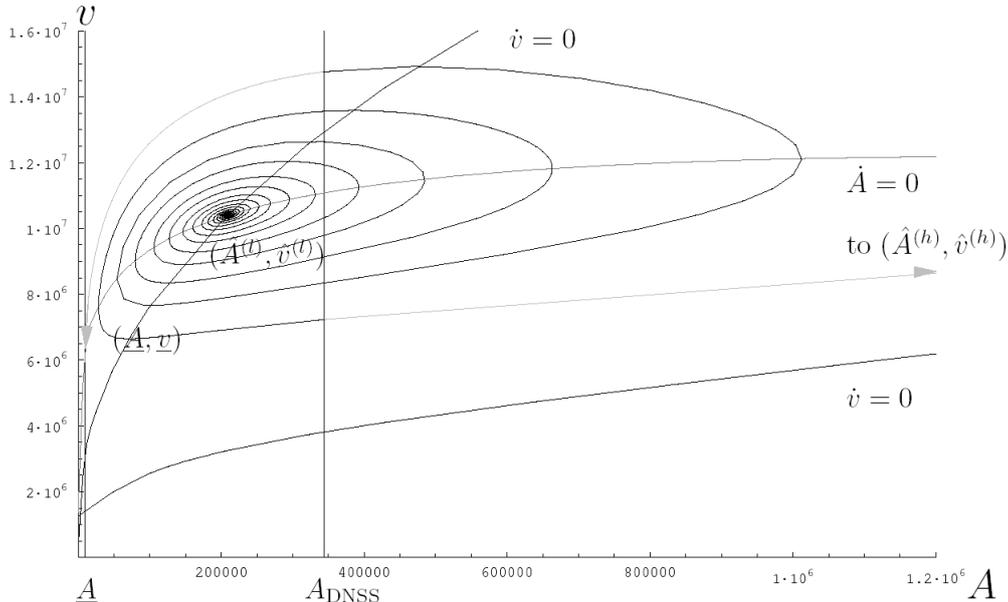
are only about 4 million children in a birth cohort in the US, that \$1B would be enough to cover every 7th grader and also half of all 4th graders with a curriculum designed for younger children.

The great advantage of prevention is that it is so inexpensive compared to treatment or incarceration. It is not extremely powerful, at least with current technology, but it is powerful enough to make it optimal to "fully fund" prevention for almost any level of the epidemic. Still, even when fully funded, prevention does not absorb a large proportion ($< 10\%$) of drug control spending.

The total optimal level of spending in equilibrium, summing across the three programs, is about \$20B per year. That is probably roughly comparable to what the US has spent historically. More precise statements are difficult to make because data are not available for *national* drug control spending *by drug*. Figures are published annually for *federal* spending to control *all drugs*. Rydell and Everingham [51] estimated that in the early 1990s, national cocaine control spending was roughly equal to federal spending on all drugs, and the federal drug control budget was \$18.8B for FY2002 (ONDCP [44]), which is quite close to the prescribed \$20B per year.⁶

⁶National budgets after 2003 have reported in a substantially different and non-

Figure 3: The DNSS threshold A_{DNSS} . The two gray curves represent the optimal policy. On the left side of the DNSS threshold, the optimal policy leads to $(\underline{A}, \underline{v})$, on the right side optimal convergence is towards $(\hat{A}^{(h)}, \hat{v}^{(h)})$.



Returning to Figure 1, in addition to the "high volume" saddle point equilibrium, there is a second "low volume" equilibrium that is an unstable focus, so the optimal policy is more complicated when control starts when the epidemic is still small. For initial numbers of users below some critical level the solution is qualitatively different than a slow approach to the high volume saddle point equilibrium.

In particular, for smaller initial numbers of users ($A(0)$) it is not possible to jump onto the stable manifold that leads to the saddle point equilibrium. If we assume there is some lower limit, \underline{A} , on the number of users (e.g., $\underline{A} = 10,000$) below which control efforts cannot drive the problem (e.g., because these residual users cannot be detected), then the point $(\underline{A}, \underline{v})$ becomes another equilibrium, where \underline{v} is given by the intersection of $A = \underline{A}$ and the isocline $\dot{A} = 0$. This steady state is approached along a trajectory which spirals out of the low volume equilibrium.

For low enough initial numbers of users it is only possible to jump on the stable manifold that approaches the lower limit equilibrium. For high enough values, it is clear one should approach the high volume equilibrium.

comparable format. Walsh [59] gives a quick, readable account of some of the changes in budgeting procedures and definitions.

For intermediate values, there is a so-called Dechert-Nishimura-Sethi-Skiba (DNSS) point (Dechert and Nishimura [20], Sethi [52],[53], Skiba [55]; cf. Grass et al. [27]) that defines two basins of attraction according to whether the optimal policy is to effectively eradicate drug use (push it to the lower limit equilibrium) or to just moderate its approach to the high volume saddle point equilibrium, as above. For the base case parameter values, that point is $A_{DNSS} = 344,339$ users.

Figure 2 shows that if the initial number of users is to the left of the DNSS point, treatment and enforcement spending are very high in absolute terms and, thus, truly enormous per user. Prevention spending is also higher than it is immediately to the right of the DNSS point, but less dramatically so. If it is optimal to eradicate the drug epidemic, then apparently it is optimal to do so aggressively and quickly (cf. Baveja et al. [1]). By spending enormous amounts on control in the early years, one avoids getting stuck at the high volume equilibrium.

Price is approximately a linear function of enforcement spending relative to market size (i.e., linear in v/A). It turns out to be a decreasing function of A for all A , with a sharp downward discontinuity at the DNSS point (since v^* is much higher just to the left of A_{DNSS} than it is just to the right of A_{DNSS}). (Figure 2.) Since when one starts to the right of A_{DNSS} one moves to the right (still assuming $A(0) < \hat{A}^{(h)}$), and when one starts to the left of A_{DNSS} one moves to the left, that means that the optimal price trajectory is very different depending on whether the optimal strategy is to eradicate or accommodate the epidemic. In particular, if the optimal strategy is to accommodate, then it is optimal to allow the price to decline over time. Enforcement spending increases, but less than proportionally in A , so v/A and, hence, p^* decreases as one approaches the high-volume saddle point equilibrium. Conversely, if the optimal strategy is to eradicate the market, then it is optimal to start with a high price and keep driving it higher and higher until A reaches its lower limit. Even though enforcement spending declines over time with the eradication strategy, A declines faster so v/A and, hence, p^* increase over time when one opts for eradication.

To summarize, at the strategic level the policy prescription is simple. When control starts, one must judge whether the current epidemic size ($A(0)$) is greater or less than the critical DNSS threshold (A_{DNSS}). If it is greater than the threshold, then the optimal strategy is to grudgingly "accommodate" the epidemic, allowing it to grow to its high-volume equilibrium ($\hat{A}^{(h)}$). Spending on all controls should increase, but less than proportionately in A so control levels increase, but control intensity decreases. If, on the other hand, the initial epidemic size is below that critical threshold, then it is optimal to "eradicate" the epidemic in the sense of pursuing all controls extremely

aggressively, quickly driving the epidemic down to its minimum size (\underline{A}).

Note: if spending were constrained to be proportional to the current size of the problem for some sufficiently small proportionality constant, e.g., because it is hard for politicians to muster support for massive spending on a problem that is currently small, then eradication might not be feasible and approaching the high-volume saddle point equilibrium might be the only alternative (cf. Tragler et al. [57]).

One final observation. The total discounted cost of the epidemic under optimal control, counting both the social costs of use and the costs of control, are monotonically increasing in the initial number of users. That is not surprising. What is surprising, is that the relationship is almost linear with a kink at the DNSS threshold. (Figure not shown.) Roughly speaking, for initial numbers of users below 1,000,000, total discounted costs increase by about \$200,000 per person increase in $A(0)$ for $A(0) < A_{DNSS}$, and by about \$80,000 per person for $A(0) > A_{DNSS}$. Those are astoundingly large numbers with a dramatic policy implication. In the absence of controls, for A near A_{DNSS} , modeled initiation is on the order of 1,000 people per day, so the cost of delaying onset of control by even a day is very large. The actual values per day of delay are not simply 1,000 times the figures above because one must account for what happens during the day of waiting. Doing so, it turns out that when the number of users is near the DNSS threshold ($A_{DNSS}/2 < A(0) < 2A_{DNSS}$), a one-day delay (or interruption) in control costs approximately \$240 million per day to the left of the DNSS threshold and \$60 million per day to its right. A corollary is that very significant investments in data collection systems may be justified if those systems can help detect future epidemics in their nascent stages.

4 Sensitivity Analysis

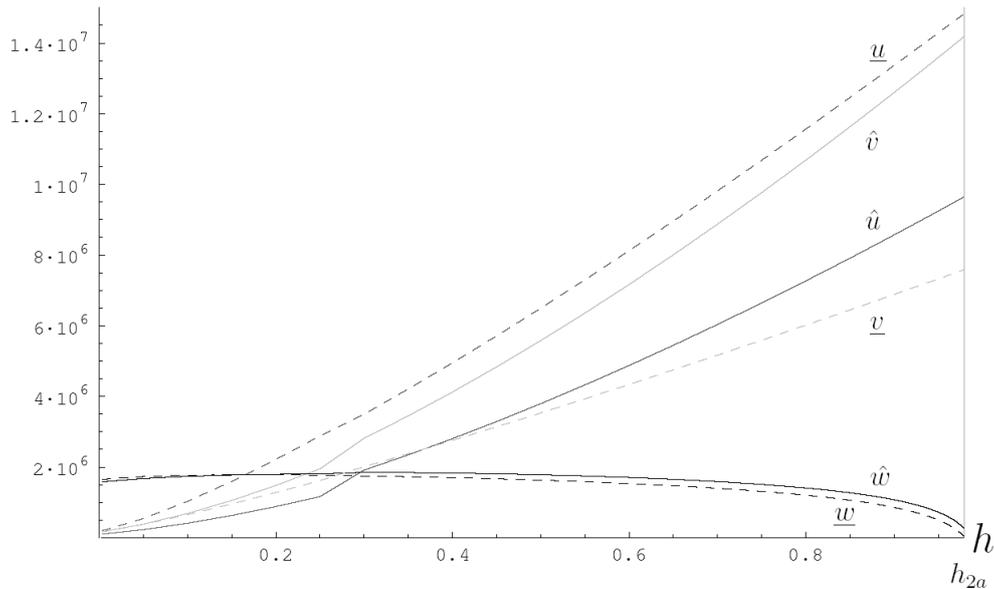
4.1 Sensitivity Analysis Concerning the Strength of Prevention

There is a reasonably strong basis for believing that current, model primary prevention technologies can reduce initiation by about $1 - h = 16\%$, but sensitivity analysis with respect to parameter h is still of interest for three reasons. First, many programs that are actually being used are not model programs, so the effectiveness of prevention today may be smaller (higher h). Second, better prevention technologies may be available in the future. For example, immunotherapies being developed to treat cocaine addiction might conceivably be used for primary prevention (Harwood and Myers [32]).

There are plausible circumstances under which such vaccinations could be highly cost-ineffective for prophylactic purposes, but the very existence of such research suggests that prevention technology is not static. Finally, a fundamental contribution of this paper is adding prevention to the mix of interventions considered, so sensitivity with respect to its performance is of particular interest.

It turns out that if more effective types of prevention were available, that could quite dramatically affect what is the optimal policy and the resulting magnitude of drug problems. Figure 4 illustrates this with regard to optimal spending on the three types of control at the lower limit (quantities with an under-bar) and the right-hand saddle equilibrium (quantities with a hat).

Figure 4: The levels of optimal control spending as functions of h at $\hat{A}^{(h)}$ (continuous) and at \underline{A} (dashed).



Moving from right to left corresponds to prevention becoming more powerful (reducing h). Not surprisingly, spending on prevention increases as it becomes more effective (until the far left when it becomes so effective that slightly reduced levels of spending are sufficient). What is striking is the extent to which spending on enforcement and treatment decline as prevention becomes more effective. Better prevention substitutes for these costly interventions. Furthermore, since prevention spending saturates at between

\$1B and and \$2B per year, total drug control spending declines as that particular drug control technology improves. Despite the declines in total control spending, with more effective prevention the right-hand saddle moves steadily to the left, roughly linearly in h , indicating fewer users in the steady state reached when accommodating the epidemic. Reduced control spending and reduced use both translate into lower social costs. Indeed, the present value of all social costs declines almost linearly by over 50% as prevention effectiveness increases enough to reduce h from 1.0 to about 0.6. That potential may justify continued investment in prevention research even though the progress to date has been more incremental than dramatic. One initially counter-intuitive result is that as prevention's effectiveness increases, the DNSS point shifts to the left, not the right. One might have expected that as the tools of drug control improved, it would be not only feasible but also desirable to eradicate epidemics even if the initial number of users were somewhat larger. However, recall that a given level of prevention spending reduces initiation by a given percentage, regardless of what that level of initiation would have been, and that initiation is increasing in the number of users. Hence, increments in prevention's effectiveness are relatively more valuable when the number of users A is large, not when it is small. Hence, while increased prevention effectiveness reduces the cost of eradicating the epidemic, it reduces the social cost from accommodating that epidemic even more, shifting to the left the DNSS point, where one is indifferent between the strategies of eradication and accommodation.

4.2 Sensitivity Analysis with Respect to Treatment Effectiveness

As mentioned, a parameter about whose value there is considerable uncertainty is the treatment effectiveness coefficient c . Our basecase value is derived from Rydell and Everingham's [51] analysis of data from the Treatment Outcomes Prospective Study, and treatment experts generally believe a 13% probability of quitting per episode of treatment is conservative. Indeed, at several points in Rydell and Everingham's analysis, they erred on the side of conservatism. Nevertheless, Manski et al. [40] note that selection effects could have introduced an upward bias and, more generally, there is next to no definitive data from randomized controlled trials concerning the effectiveness of cocaine treatment. Hence, this parameter is an appropriate object of sensitivity analysis.

Varying this parameter affects the saddle-point equilibrium in predictable ways. The more effective treatment is, the greater its share of control spend-

ing in steady state, and the fewer users there are in steady state. In particular, if treatment were 1% more effective, it would be optimal in steady state to spend about 1% more on treatment and almost 1% less on enforcement (+0.97% and -0.86% , respectively, to be precise). Even though enforcement spending declines, enforcement intensity increases because the decline in the number of users is even greater (-1.65%), inflating the ratio of v over A . Prevention spending also declines but less dramatically (by 0.22%), which is consistent with the general finding that the optimal level of prevention spending is stable in multiple respects. Overall, improved treatment technology acts as a substitute for enforcement and prevention. Indeed, because with base case parameter values more is spent on enforcement (\$11.4B) than treatment (\$7.8B), the increase in treatment effectiveness actually leads to a reduction in total steady-state control spending.

4.3 Sensitivity Analysis with Respect to Initiation Exponent α

It is generally presumed that initiation is an increasing but concave function of the current number of users, modeled here as initiation being proportional to the current number of users A raised to an exponent α , with $\alpha = 0.3$ in the base case. Sensitivity analysis with respect to α is of interest because prevention is related to initiation and because it turns out that the location of the DNSS point is greatly affected by the value of α .

When α is varied, we vary k as well to keep the rate of initiation under base case conditions constant at 1,000,000 per year. That means that as α is reduced, the leading coefficient k is increased, and rather dramatically. By definition the reduction in α exactly offsets the increase in k when the number of users is 6.5 million, but for smaller numbers of users typical of earlier stages of the epidemic, the increase in k dominates. So in these sensitivity analyses, reducing α implies increasing rather substantially the force or "power" of initiation early in the epidemic.

Predictably, then, reducing α moves the DNSS point to the left, implying that eradication is the optimal strategy only under narrower circumstances. That makes sense. The appeal of eradication is that one drives use down to such a low level that initiation is also modest. When α is smaller, initiation with small A is much greater, so the benefit from reduced initiation achieved by driving A down to \underline{A} is much smaller.

Still, the extent to which this turns out to be the case is striking. If α drops merely to 0.28, the DNSS point disappears and accommodation is always optimal. On the other hand, if α increases to 0.3415, the DNSS point

moves so far to the right that it reaches the high-value saddle equilibrium (which also has been moving left), implying that eradication is always the optimal strategy.

4.4 Sensitivity Analysis Concerning the Social Cost per Gram Consumed

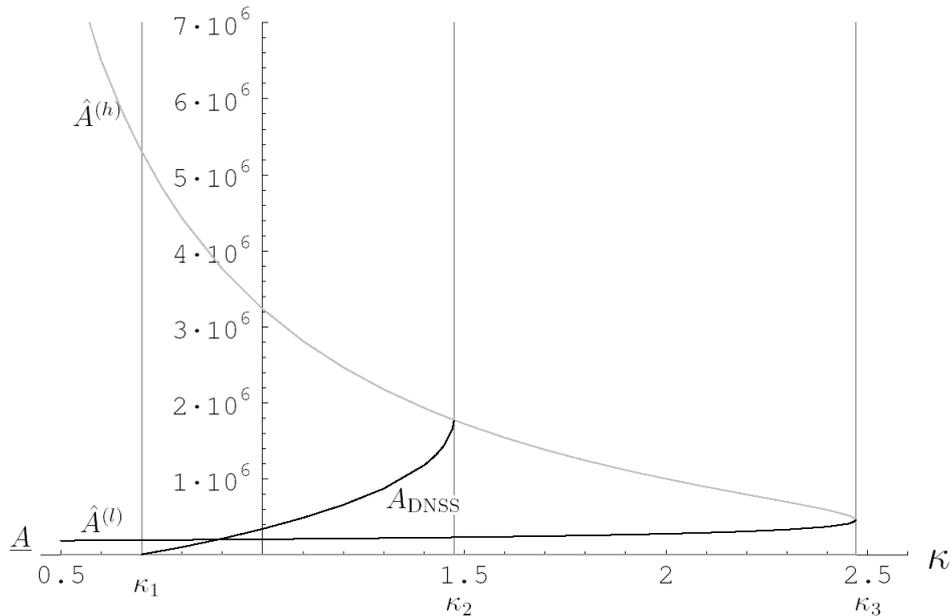
We observed that the optimal total level of spending at the saddle point equilibrium may be roughly comparable to what the US has spent historically on cocaine control. However, what level is optimal depends substantially on the presumed social cost per gram of cocaine consumed, and there is considerable uncertainty as to whether the base case value ($\kappa = \$100/\text{gram}$) is "correct", both because of data limitations and because there can be genuine disagreement concerning what categories of costs should be included as social costs.⁷ Generally, the greater the perceived social cost per unit of consumption, the more it is optimal to spend at the saddle point equilibrium and, hence, the lower the level of use in that equilibrium. In particular, if the social cost per gram were believed to be 20% higher, then the optimal level of drug control spending at the saddle equilibrium would be 11% higher. Likewise, if κ were 20% lower, the optimal steady state spending would be 17% lower, with the changes being most dramatic for treatment and least dramatic for prevention.

In contrast, the level of control spending at the lower limit \underline{A} is almost unaffected by κ , presumably because the value of that spending is whatever it takes to prevent an epidemic from exploding, not an amount that is determined by balancing current control costs with current social costs of use.

Sensitivity of the optimal policy to variation in the assumed social cost per gram of use is even more pronounced for larger variations from the base case. In particular, reducing κ affects the DNSS threshold in qualitatively the same way as reducing α does, as is illustrated in Figure 5, albeit for quite different reasons. As κ declines, the DNSS threshold shifts to the left, disappearing when κ drops to 0.7. Similarly, the DNSS threshold shifts to the right as κ increases, merging with the saddle point equilibrium when $\kappa = 1.474$.

⁷Notable examples include social costs borne by family members, any benefits of use of an illicit substance, valuation of a human life beyond that person's labor market earnings, and valuation of pain and suffering associated with crime and with addiction itself.

Figure 5: The influence of κ on the equilibrium values and the DNSS threshold. The relation between κ and the high equilibrium $\hat{A}^{(h)}$ is displayed in the upper gray branch, while the black lower branch shows the relation between κ and the unstable focus at $\hat{A}^{(l)}$. The black curve between κ_1 and κ_2 bending upwards represents the level of the DNSS threshold, and the horizontal gray line at the very bottom stands for the lower limit at \underline{A} .



Hence, someone who thinks the social costs per gram of cocaine use are less than \$70 per gram ought always to favor accommodation, whereas someone who thinks they are over \$147 per gram ought always to strive for eradication, even if the epidemic is already established. That is striking sensitivity inasmuch as it is easy for two reasonable people to disagree by a factor of 2 or more concerning the social cost per gram of cocaine.

An obvious implication is a plausible explanation for the persistent heated disagreements between drug policy "hawks" who favor having the goal be a "drug-free America" and "doves" who think the social costs of eradication exceed its benefits.

A more subtle point emerges from the observation that the social cost per gram consumed is not an immutable physical constant like π or the speed of light. There are a whole set of policies not modeled here but popular in countries such as Australia and the Netherlands that go under the banner of "harm reduction". That term is highly controversial and widely misused and misunderstood. For the moment, let it mean simply and literally programs

that reduce the social harm per unit of drug used, i.e., that reduce κ . The paradigmatic harm reduction policy, distributing clean syringes to injection drug users, is largely irrelevant for cocaine in the US, which is not primarily injected. Another favorite of harm reduction advocates is increasing treatment availability, which is already included in the model and is not actually likely to have as its primary outcome reductions in κ . Still, one can imagine other harm reduction tactics that would be relevant for cocaine in the US, including offering various forms of social support to the families of cocaine abusers, particularly their children; developing immunotherapies that treat cocaine overdose more effectively (Harwood and Myers [32]); and pursuing different types of law enforcement that push street markets into forms of distribution that generate less violence per kilogram sold and used, rather than seeking to reduce use by driving up prices.⁸ Whatever the specifics, according to this model there can be a strong interaction between the presence of effective harm reduction and whether the optimal policy is eradication or accommodation. If one can design harm reduction strategies that reduce the average social cost per gram consumed, then accommodation might be the better alternative, even if eradication would be preferred in their absence.

4.5 Sensitivity Analysis Concerning the Lower Limit on the Number of Users

The larger the lower limit, \underline{A} , below which control cannot drive the number of users, the smaller the DNSS point. For example, doubling \underline{A} from 10,000 to 20,000 roughly reduces the DNSS point by two thirds (reduces it from 334,339 to 128,268). This seemingly counter-intuitive result has a simple explanation. The smaller the lower limit on A , the more appealing that low-volume steady state is and, hence, the more the decision maker would be willing to invest in order to drive the epidemic to that lower steady state. Willingness to invest more means being willing to pursue the "eradication" strategy even if the initial number of users is somewhat larger.

If the minimum number of users is interpreted as the number below which users are essentially invisible, this has an interesting implication. Policy makers would like to push that lower limit down as far as possible. Doing so raises the DNSS point and, thus, increases the time it takes an epidemic to reach the "point of no return", beyond which the best that policy can do is

⁸One partial explanation for why homicides have fallen so dramatically in New York City may be that much retail drug distribution has shifted from anonymous street markets where controlling "turf" produces profits to instances in which seller-user dyads arrange private meetings in covert locations, often using cell phones.

moderate expansion to the high volume equilibrium.

As noted above, similar logic explains the otherwise surprising result that the more effective prevention is (i.e., the lower h is) the lower is the DNSS threshold.

5 Discussion

The analysis here confirms the observation of Behrens et al. [2] and Tragler et al. [57] that it can be misleading to discuss the merits of different drug control interventions in static terms (e.g., asserting that prevention is better than enforcement or vice versa without reference to the stage of the epidemic). Even this simple model of drug use and drug control can yield optimal solutions that involve substantially varying the mix of interventions over time.

Furthermore, the broad outlines of the policy recommendations are similar to those in Tragler et al. [57]. When a new drug problem emerges, policy makers must choose whether to essentially eradicate use or to accommodate the drug by grudgingly allowing it to grow toward a high-volume equilibrium. If the decision is to eradicate, then control should be very aggressive, using truly massive levels of both enforcement and treatment relative to the number of users to drive prevalence down as quickly as possible. If accommodation is pursued, levels of spending on price-raising enforcement, treatment, and primary prevention should increase linearly but less than proportionally with the number of users (i.e., linearly with a positive intercept). So the total level of drug control spending should grow as the epidemic matures, but spending per user would decline.

Of all the interventions, optimal spending on primary prevention is least dependent on the stage of the epidemic. To a first-order approximation, prevention spending should be about the same throughout. With our particular parameterization, that level is roughly enough to offer a good school-based program to every child in a birth cohort, but not dramatically more than that. That relative independence on the state of the epidemic is fortuitous inasmuch as there are built in lags to primary prevention, at least for school-based programs. Such programs are usually run with youth in junior high, but the median age of cocaine initiation in the US is 21 (Caulkins [4]).

However, these observations do not in any way imply that adding prevention to this dynamic model does not alter the results. Prevention is a strong substitute for price-raising enforcement and treatment. The more effective prevention is, the less that should be spent on those other interventions. Furthermore, a truly effective prevention program would be such a strong

substitute that both the amount of drug use and the combined optimal levels of drug control spending would decline, leading of course to a substantial reduction in the total social costs associated with the drug epidemic.

The catch is that to date even the better primary prevention programs seem to be only moderately effective (Caulkins et al. [13],[14]), and the programs actually implemented are often not the best available (Hallfors and Godette [30]). Hence, with respect to the wisdom of further investments in improving the "technology" of primary prevention, one can see the glass as half full or half empty. The pessimists would point to limited progress to date and suggest focusing elsewhere. The optimists would see the tremendous benefits that a truly effective primary prevention program would bring and redouble their efforts.

The second broad policy contribution of this paper relative to the prior literature is the sensitivity analysis with respect to the location of the DNSS threshold and, hence, of when each broad strategy (eradication or accommodation) is preferred. In short, the finding is that the location of the DNSS threshold is highly sensitive to three quantities that are difficult to pin down for various reasons: the social cost per gram of cocaine consumed, the exponent in the initiation function governing how contagious the spread of drug use is, and the lower limit on prevalence below which it is assumed that control cannot drive the epidemic.

A depressing implication is that it will generally be exceedingly difficult to make an informed decision concerning the strategic direction for policy concerning a newly emergent drug. More is known and more data are available about the current cocaine epidemic in the US than about any other epidemic of illicit drug use, yet these parameters still cannot be pinned down even for cocaine in the US. It is hard to imagine that when a new drug epidemic emerges, we will have better information about it, at least at that early stage, and one of the results above was a startlingly high increase in social cost for each *day* that initiation of control is delayed. So a "wait and study" option may not be constructive.

Another depressing implication concerns the result for the lower limit on prevalence and its interpretation in a world of polydrug use. The model considered explicitly just one drug, cocaine. If there were just one illicit drug entering a "virgin" population, it might be somewhat plausible to drive use of that drug down to very low levels. However, the US already has several million dependent drug users who tend to use a wide variety of drugs, including new ones that come along. So if the US now faced a new epidemic, it might be that the only way it could drive use of that drug down to levels such as the lower limit considered here, would be to also eliminate use of the existing established drugs such as cocaine, heroin, and methamphetamine. That

may be impossible or at least, according to this model, likely not optimal. Inasmuch as higher lower limits on prevalence make eradication strategies less appealing, accommodation may be the best option for future epidemics, even if eradication would have been the better course if we could turn back the clock to 1965.

The one positive observation, though, is that there exist, at least in theory, another set of drug control interventions, not modeled here, that would target not drug use but the objective function coefficient associated with that use. Introducing interventions of that sort into this framework would be one of many productive avenues for further research.

Acknowledgements

This research was financed in part by the Austrian Science Fund (FWF) under grant P25979-N25. We thank Gustav Feichtinger and Florian Moyzisch for their contributions to this paper.

Appendix: Optimality Conditions

The current value Hamiltonian H is given by

$$H = -(\kappa\theta Ap^{-\omega} + u + v + w) + \lambda(kA^\alpha p^{-a}\Psi - c\beta A - \mu p^b A),$$

where λ describes the current-value costate variable.

Note that it is not necessary to formulate the maximum principle for the Lagrangean, which incorporates the non-negativity constraints for the controls, since u , v , and w all turn out to be positive in the analysis described in this paper.

According to Pontryagin's maximum principle we have the following three necessary optimality conditions:

$$u = \arg \max_u H,$$

$$v = \arg \max_v H,$$

and

$$w = \arg \max_w H.$$

Due to the concavity of the Hamiltonian H with respect to (u, v, w) , setting the first order partial derivatives equal to zero leads to the unrestricted extremum, and we get the following expressions for the costate λ :

$$H_u = 0 \Rightarrow \lambda = \frac{-1}{c\beta_u A}, \quad (1)$$

$$H_w = 0 \Rightarrow \lambda = \frac{1}{kp^{-a} A^\alpha \Psi_w}, \quad (2)$$

$$H_v = 0 \Rightarrow \lambda = \frac{1 - \kappa\theta\omega p^{-\omega-1} p_v A}{-akp^{-a-1} p_v A^\alpha \Psi - \mu b p^{b-1} p_v A}, \quad (3)$$

where subscripts denote derivatives w.r.t. the corresponding variables.

The concavity of the maximized Hamiltonian with respect to the state variable, however, cannot be guaranteed, so the usual sufficiency conditions are not satisfied.

With Equations (1)-(3) we can describe u , w , and λ as functions of A and v as follows:

$$\lambda(A, v) := \frac{\frac{p_v}{p} \left(\frac{a}{m} + \kappa\theta\omega p^{-\omega} A \right) - 1}{ahkp^{-a-1} p_v A^\alpha + \mu b p^{b-1} p_v A}, \quad (4)$$

$$u(A, v) := \left(\frac{-(A + \delta)^z}{czA\lambda(A, v)} \right)^{\frac{1}{z-1}},$$

$$w(A, v) := \frac{1}{m} \ln \left((h-1) k m p^{-a} A^\alpha \lambda(A, v) \right).$$

Due to this simplification we can concentrate on the two variables A and v .

To gain an equation for \dot{v} we differentiate $\lambda(A, v)$ with respect to time:

$$\dot{\lambda} = \lambda_A \dot{A} + \lambda_v \dot{v}. \quad (5)$$

Setting (5) equal to the costate equation

$$\dot{\lambda} = r\lambda - H_A,$$

yields:

$$\dot{v} = \frac{r\lambda - H_A - \lambda_A \dot{A}}{\lambda_v},$$

where we insert $\lambda(A, v)$ from (4) and the corresponding derivatives λ_A and λ_v as well as H_A given by

$$H_A = -\kappa\theta p^{-\omega-1} (p - \omega p_A A) + \lambda [kp^{-a-1} \Psi(\alpha A^{\alpha-1} p - a A^\alpha p_A) - c(\beta_A A + \beta) - \mu p^{b-1} (b p_A A + p)].$$

References

- [1] Baveja, A., Caulkins, J.P., Liu, W., Batta, R., Karwan, M.H.: When haste makes sense: Cracking down on street markets for illicit drugs. *Socio-Economic Planning Sciences* **31**, 293-306 (1997)
- [2] Behrens, D.A., Caulkins, J.P., Tragler, G., Haunschmied, J., Feichtinger, G.: Optimal control of drug epidemics: Prevent and treat – but not at the same time. *Management Science* **46**(3), 333-347 (2000)
- [3] Caulkins, J.P.: The cost-effectiveness of civil remedies: The case of drug control interventions. In Green Mazerolle, L., Roehl, J. (eds.): *Crime Prevention Studies* **9**, 219-237 (1998a)
- [4] Caulkins, J.P.: Drug prevention: The paradox of timing. *Federation of American Scientists' Drug Policy Analysis Bulletin* **5**, 1-3 (1998b)
- [5] Caulkins, J.P.: The dynamic character of drug problems. *Bulletin of Narcotics* **53**(1), 11-23 (2001)
- [6] Caulkins, J.P. Drug policy: Insights from mathematical analysis. In Brandeau, M.L., Sainfort, F., Pierskalla, W.P. (eds.): *Operations Research and Healthcare: A Handbook of Methods and Applications*, 297-332. Kluwer Academic Publishers, Boston (2004)
- [7] Caulkins, J.P.: Models pertaining to how drug policy should vary over the course of a drug epidemic. *Substance Use: Individual Behavior, Social Interactions, Markets, and Politics. Advances in Health Economics and Health Services Research* **16**, 407-439 (2005)
- [8] Caulkins, J.P., Behrens, D.A., Knoll, C., Tragler, G., Zuba, D.: Modeling dynamic trajectories of initiation and demand: The case of the US cocaine epidemic. *Health Care Management Science* **7**(4), 319-329 (2004)
- [9] Caulkins, J.P., Chiesa, J., Everingham, S.S.: Response to the National Research Council's Assessment of RAND's Controlling Cocaine Study. MR-1265, RAND, Santa Monica, CA (2000)
- [10] Caulkins, J.P., Reuter, P.: What price data tell us about drug markets. *Journal of Drug Issues* **28**(3), 593-612 (1998)
- [11] Caulkins, J.P., Reuter, P.: How drug enforcement affects drug prices. In Tonry, M. (ed.): *Crime and Justice - A Review of Research* **39**, University of Chicago Press, Chicago, 213-272 (2010)

- [12] Caulkins, J.P., Rydell, C.P., Schwabe, W.L., Chiesa, J.: Mandatory Minimum Drug Sentences: Throwing Away the Key or the Taxpayers' Money? RAND, Santa Monica, CA (1997)
- [13] Caulkins, J.P., Paddock, S., Pacula, R., Chiesa, J.: School-based drug prevention: What kind of drug use does it prevent? RAND, Santa Monica, CA (2002)
- [14] Caulkins, J.P., Rydell, C.P., Everingham, S.S., Chiesa, J., Bushway, S.: An Ounce of Prevention, a Pound of Uncertainty: The Cost-Effectiveness of School-Based Drug Prevention Program. RAND, Santa Monica, CA (1999)
- [15] Chaloupka, F.J., Grossman, M., Tauras, J.A.: The demand for cocaine and marijuana by youth. In Chaloupka, F.J., Grossman, M., Bickel, W.K., Saffer, H. (eds.): *The Economic Analysis of Substance Use and Abuse: An Integration of Econometric and Behavioral Economic Research*, 133-155. University of Chicago Press, Chicago (1999)
- [16] Chaloupka, F.J., Pacula, R.L.: Economics and anti-health behavior: The economic analysis of substance use and abuse. In Bickel, W., Vuchinich, R. (eds.): *Reframing Health Behavior Change with Behavioral Economics*, 89-111. Lawrence Earlbaum Associates, Hillsdale, NJ (2000)
- [17] Crane, B.D., Rivolo, A.R., Comfort, G.C.: An Empirical Examination of Counterdrug Program Effectiveness. IDA Paper P-3219, Institute for Defense Analysis, Alexandria, VA (1997)
- [18] Cuijpers, P.: Three decades of drug prevention research. *Drugs: Education, Prevention and Policy* **10**(1), 7-20 (2003)
- [19] Dave, D.: The Effects of Cocaine and Heroin Price on Drug-Related Emergency Department Visits. NBER Working Paper, Cambridge, MA (2004)
- [20] Dechert, W.D., Nishimura, K.: 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 (1983)
- [21] DeSimone, J.: The Effect of Cocaine and Heroin Prices and Arrests on Cocaine and Heroin-Related Deaths. Working Paper (2001)
- [22] DeSimone, J., Farrelly, M.C.: Price and enforcement effects on cocaine and marijuana demand. *Economic Inquiry* **41**(1), 98-115 (2003)

- [23] Feichtinger, G., Hartl, R.F.: Optimale Kontrolle Ökonomischer Prozesse – Anwendungen des Maximumprinzips in den Wirtschaftswissenschaften. DeGruyter, Berlin (1986)
- [24] Ferrence, R.: Diffusion theory and drug use. *Addiction* **96**, 165-173 (2001)
- [25] Gallet, C.A.: Can price get the monkey off our back? A meta-analysis of illicit drug demand. *Health Economics* **23**, 55-68 (2014)
- [26] Golub, A., Johnson, B.D. The crack epidemic: Empirical findings support a hypothesized diffusion of innovation process. *Socio-Economic Planning Sciences* **30**(3), 221-231 (1996)
- [27] Grass, D., Caulkins, J.P., Feichtinger, G., Tragler, G., Behrens, D.A.: Optimal Control of Nonlinear Processes: With Applications in Drugs, Corruption and Terror. Springer, Berlin (2008)
- [28] Greenwood, P.W., Rydell, C.P., Abrahamse, A.F., Caulkins, J.P., Chiesa, J.R., Model, K.E., Klein, S.P.: Three Strikes and You're Out: Estimated Benefits and Costs of California's New Mandatory-Sentencing Law. MR-509-RC, RAND, Santa Monica, CA (1994)
- [29] Grossman, M., Chaloupka, F.J.: The demand for cocaine by young adults: A rational addiction approach. *Journal of Health Economics* **17**, 427-474 (1998)
- [30] Hallfors, D., Godette, D.: Will the 'Principles of Effectiveness' improve prevention practice? Early findings from a diffusion study. *Health Education Research* **17**(4), 461-470 (2002)
- [31] Harwood, H., Fountain, D., Livermore, G.: The Economic Costs of Alcohol and Drug Abuse in the United States, 1992. National Institute on Drug Abuse and National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD (1998)
- [32] Harwood, H.J., Myers, T.G. (eds.): New Treatments for Addictions. National Academy Press, Washington, DC (2004)
- [33] Institute of Medicine (IOM): Treating Drug Problems. National Academy Press, Washington, DC (1990)
- [34] Institute of Medicine (IOM): Pathways of Addiction: Opportunities in Drug Abuse Research. National Academy Press, Washington, DC (1996)

- [35] Kaplan, J.: Heroin: The Hardest Drug. University of Chicago Press, Chicago (1983)
- [36] Kleiman, M.A.R.: Crackdowns: The effects of intensive enforcement on retail heroin dealing. In Chaiken, M.R. (ed.): Street-Level Drug Enforcement: Examining the Issues. National Institute of Justice, Washington, DC (1988)
- [37] Kleiman, M.A.R.: Enforcement swamping: A positive-feedback mechanism in rates of illicit activity. *Mathematical and Computer Modeling* **17**, 65-75 (1993)
- [38] Lonard, D., Long, N.V.: Optimal Control Theory and Static Optimization in Economics. Cambridge University Press, Cambridge, MA (1992)
- [39] Manski, C.F., Pepper, J.V., Petrie, C.V. (eds.): Assessment of Two Cost-Effectiveness Studies on Cocaine Control Policy. National Academy Press, Washington, DC (1999)
- [40] Manski, C., Pepper, J., Petrie, C.: Informing America's Policy on Illegal Drugs: What We Don't Know Keeps Hurting Us. National Academy Press, Washington, DC (2001)
- [41] Miller, T.R., Cohen, M.A., Wiersema, B.: Victim Costs and Consequences: A New Look. National Institute of Justice, Washington, DC (1996)
- [42] Moore, M.H.: Achieving discrimination on the effective price of heroin. *American Economic Review* **63**, 270-277 (1973)
- [43] Office of National Drug Control Policy (ONDCP): The National Drug Control Strategy, 1996. The White House, Washington, DC (1996)
- [44] Office of National Drug Control Policy (ONDCP): The National Drug Control Strategy, 2002. The White House, Washington, DC (2002)
- [45] Reuter, P.: Disorganized Crime: The Economics of the Visible Hand. MIT Press, Cambridge, MA (1983)
- [46] Reuter, P., Kleiman, M.A.R.: Risks and prices: An economic analysis of drug enforcement. In Morris, N., Tonry, M. (eds.): Crime and Justice: A Review of Research. University of Chicago Press, Chicago, IL (1986)
- [47] Rhodes, W., Johnston, P., Han, S., McMullen, Q., Hozik, L.: Illicit Drugs: Price Elasticity of Demand and Supply. ABT Associates (2001)

- [48] Riley, K.J.: Crack, Powder Cocaine, and Heroin: Drug Purchase and Use Patterns in Six U.S. Cities. National Institute of Justice, Washington, DC (1997)
- [49] Rocheleau, A.M., Boyum, D.: Measuring Heroin Availability in Three Cities. Office of National Drug Control Policy, Washington, DC (1994)
- [50] Rydell, C.P., Caulkins, J.P., Everingham, S.S.: Enforcement or treatment? Modeling the relative efficacy of alternatives for controlling cocaine. *Operations Research* **44**, 1-9 (1996)
- [51] Rydell, C.P., Everingham, S.S.: Controlling Cocaine. Supply Versus Demand Programs. RAND, Santa Monica, CA (1994)
- [52] Sethi, S.P.: Nearest feasible paths in optimal control problems: Theory, examples, and counterexamples. *Journal of Optimization Theory and Applications* **23**(4), 563-579 (1977)
- [53] Sethi, S.P.: Optimal advertising policy with the contagion model. *Journal of Optimization Theory and Applications* **29**(4), 615-627 (1979)
- [54] Sevigny, E.L., Caulkins, J.P.: Kingpins or mules? An analysis of drug offenders incarcerated in federal and state prisons. *Criminology and Public Policy* **3**(3), 401-434 (2004)
- [55] Skiba, A.K.: Optional growth with a convex-concave production function. *Econometrica* **46**, 527-539 (1978)
- [56] Tragler, G., Caulkins, J.P., Feichtinger, G. The Impact of Enforcement and Treatment on Illicit Drug Consumption. Working Paper 212, Department of Operations Research and Systems Theory, Vienna University of Technology, Vienna (1997)
- [57] Tragler, G., Caulkins, J.P., Feichtinger, G.: Optimal dynamic allocation of treatment and enforcement in illicit drug control. *Operations Research* **49**(3), 352-362 (2001)
- [58] United Nations Office on Drugs and Crime (UNODC): 2004 World Drug Report. Oxford University Press (2004)
- [59] Walsh, J.M. Fuzzy Math: Why the White House Drug Control Budget Doesn't Add Up. *FAS Drug Policy Analysis Bulletin* **10** (2004)