How to learn the hard way

DORIS BEHRENS –GUSTAV FEICHTINGER – GERNOT TRAGLER

DEPARTMENT OF OPERATIONS RESEARCH AND SYSTEMS THEORY
VIENNA UNIVERSITY OF TECHNOLOGY

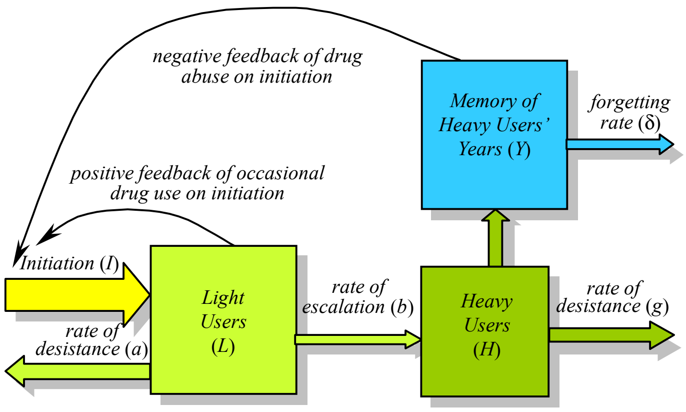

Figure 1: Flow-diagram of the *LHY*-model

Are you one of those regarding tax evasion as a very noble form of sports? If so you are exposing yourself to the increased risk of criminal prosecution by fiscal authorities and – in case of detection – serving as a deterrent paradigm for this kind of behavior.

Likewise with "sins of omission" in road traffic such as "forgetting" to put on the safety belt or the crash helmet. Observing an accident – possibly even with fatal consequences – will remind us of that, at least for a little while.

Let us finally think about consuming an (il)legal substance such as nicotine, alcohol or cocaine. Quite often it is just a small step to escalated from occasional use to addiction. Physical, psychological, social and economic suffering associated with abuse, however, gives clear evidence of the considerable risk already associated with occasional use.

What is common to these examples from daily life, even if not necessarily your own?

1. The basic mechanism

We can summarize the basic features of the potentially risky behavior sketched by the above examples as follows (see also Figure 1):

- *Heterogeneity*
  Modeling the entire spectrum of possible human behavior is certainly desirable but almost always illusory – at least for applications based on data from real world problems. Hence, the "*LHY* model" offers an alternative procedure for approaching modeled heterogeneity by focusing on a simple distinction between two stages (or intensities) of behavior, where the first stage (L) covers behavior not necessarily illegal but bearing the risk of escalating to harmful or noxious behavior (H). (For the mathematical derivation of the *LHY* model consult Behrens et al. (2000b).)

For a better understanding let us think about heterogeneity in terms of the examples mentioned above. "Stage 1 individuals" include among others authors of "creative" tax returns, "social smokers", "occasional drinkers", etc. and all those who consider that safety belts and helmets have got nothing to do with their own safety in road traffic. The corresponding road casualties, addicts and convicted tax evaders, however, belong to the set of "stage 2 individuals" (also see Table 1).

Table 1: Examples for the *LHY*-model's possible fields of application

Problem-type	**Stage 1 (risky)**	**Stage 2 (harmful)**	**Control interventions**
"taxes "	Tax evasion	Detected	Tax audits
"seat-belts"	Don't buckle up	Road casualties	Information campaigns, controls
"illicit drugs"	Light use	Addictive (heavy) use	Prevention, treatment, enforcement

Certainly this simple dichotomous distinction between two stages of behavior might appear as an over-simplification at first sight but note that these models – as described in the following sections – beautifully reflect quite a large number of real-world problems. Additionally, they do not reject econometric estimations as more-stage models mainly would do.

- *Escalation*
  Per time unit a fixed fraction (b) of "stage 1 individuals" escalates to become "stage 2 individuals". In terms of our examples escalation is achieved by tax audits, by traffic accidents, or by becoming a heavy (=addicted) drug user.

- *Deterrence and Infection*
  According to the motto "If he has / can / does / risks that, then so will I" the stage 1 behavior spreads like an annoying flu on cold and wet autumn days. A second (also endogenous) mechanism counteracts this spreading, however. The fact that the stage 2 behavior has – using a little euphemism – unpleasant side-effects dampens this contagious effect as opposed to a flu. Deterrence is not solely based on the physical existence of stage 2 individuals but also on the memory of past adverse experiences and on the information about the consequences of harmful stage 2 behavior (represented by the state Y).

  Hence, modeled initiation into potentially risky behavior (I) is non-linearly moderated by the ratio of stage 1 to stage 2 individuals (H/L and Y/L, respectively). More, precisely, the rate at which current stage 1 individuals "recruit" initiates is moderated by the "reputation" or image the type of behavior has, and the reputation is governed by H/L and Y/L, respectively, not the absolute number of stage 2 individuals. In terms of the "drug-example" we might think of the following situation: Even if there were a number of heavy users, the drug might appear benign if they were buried in a sea of (relatively happy) light users or if potential recruits have no information (Y) about them. This might be caused by a fast forgetting (δ) process or by insufficient information campaigns (q).

Desistance can happen instantaneously (a) or "later on" from the second stage of behavior, H, caused by treatment, limitation, recovery or death (g).

The *LHY* model allows to gain insight about the development of potentially risky behavior over time. A decision-maker, however, might be at least as interested in the cost-effectiveness analysis as in the underlying descriptive model. Therefore it is necessary to weigh out the discounted stream of the costs of control and the loss caused by L and H. As far as the safety belt is concerned the control instrument may consist of traffic checks or campaigns, the costs of which will have to be weighed up with those resulting from an accident (for other examples see "control interventions" in Table 1).

The following section summarizes an application of the econometrically estimated *LHY* model used as constraint for an optimal control problem where the objective is determined as described above.

2. U.S. Cocaine

Modeling the "drug problem", especially optimal demand-side drug control interventions such as prevention and treatment of addicted users (see, e.g., Behrens et al., 2000a, 2002; compare Caulkins et al., 2000; Tragler et al., 2001), is based upon a Markovian model of population flows where initiation into drug use is scripted and which was originally introduced by Susan S. Everingham and C. Peter Rydell (1994). Everingham and Rydell suggested that, at least for cocaine, a simple dichotomous distinction between "light" (L) and "heavy" (H) users is sufficient.

Behrens et al. (1999) converted this model of drug demand into continuous time and incorporated a feedback effect of the current prevalence, or level, of use on initiation into new use. In particular, they included an endogenous model of initiation in which some initiation was "spontaneous" (e.g., because of immigration) but most occurred through interactions with current light users because most people who start using drugs do so through contact with a friend or sibling who is already using. Indeed, the metaphor of a drug "epidemic" is commonly used precisely because of this tendency for current users to "recruit" new users (see Figure 2). If that were the only mechanism by which current use affected initiation one might expect initiation to increase monotonically. David F. Musto (1987) has argued that, in addition, knowledge of the possible adverse effects of drug use acts as a deterrent or brake on initiation. He hypothesizes that drug epidemics eventually die out when a new generation of potential users becomes aware of the dangers of drug abuse and, as a result, does not start to use drugs. Whereas many light users work, uphold family responsibilities, and generally do not manifest obvious adverse effects of drug use, a significant fraction of heavy users are visible reminders of the dangers of using addictive substances. Hence, one might expect large numbers of heavy users to suppress rates of initiation into drug use.

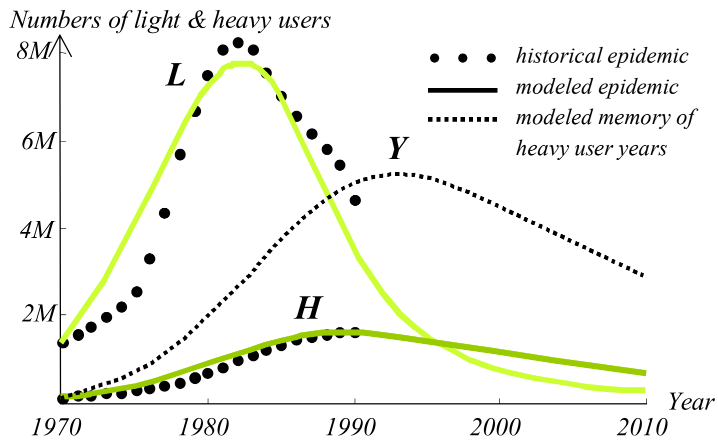

Figure 2: Plot of the historically observed (Everingham & Rydell) and the modeled cocaine epidemic currently observed in the US including the memory of abuse (for a=0.163, b=0.024, g=0.062, s=0.61, τ=50,000, δ=0.291, q=3.443; L_0=1,400,000, H_0=130,000, Y_0=110,000).

The specification of the parameter values is done for the cocaine epidemic currently observed in the United States, because of its magnitude and because data on that epidemic are reliable. The flow rates a and b in the *LHY* model are computed as the time-continuous equivalents of the Everingham-Rydell estimates (Everingham and Rydell, 1994, p.43). For example, based on the *National Institute on Drug Abuse* (NIDA, 1991) they estimate that 15% of light users quit use each year and 2.4% escalate to heavy use, so according to Behrens et al (1999) we set $a \cong 0.163$ and $b \cong 0.024$. Additionally, Everingham and Rydell (1994, p.42) gave ranges of parameters for which they found good fits to historical data reported by the National Institute on Drug Abuse (NIDA, 1991), which are $a \in (0.157, 0.168)$ and $b \in (0.02, 0.03)$. For reasons described in Behrens et al. (1999) we choose $g \cong 0.062$. [1]

Initiation (I) depends on light users and the decaying memory of heavy-user years. Hence, we have to estimate values for the "initiation parameters", the number of innovators, τ, the raw infection rate, s, and the measure of the deterrent effect of Y/L, q. For the determination of the parameters s and τ we follow Behrens et al. (1999) yielding the values $s = 0.61$ and $\tau = 50{,}000$. The corresponding size of parameter s can be interpreted as indicating that approximately two light users would "persuade" one non-user per year to try cocaine, if there was no memory of years of abuse giving cocaine a bad reputation.

The constant q measures the deterrence caused by the memory of heavy user years[2] and is, consequently, correlated with the "forgetting rate", δ.[3] The values for δ and q, respectively, are found by minimizing of the sum of the squared difference between the modeled and observed annual cocaine initiation rates from 1971 to the year x, where x varies from 1980–1989 (with initial numbers of users set equal to the number Everingham and Rydell (1994) estimate were present in 1970, where each of the 10 estimation runs was performed for (a) $Y_0 = H_0$, and (b) $Y_0 = 0.5*H_0/\delta$). For the tables including the appropriate pairs of parameter estimates, the LS values of deviation, and a visualization of the results in graphical form see Knoll and Zuba (2000).

Before we focus on the normative models one final, interesting characteristic of this model is worth noting. Large values of Y_0 support (speed up) the burning out of the drug epidemic of its own accord. This might help to explain why Europe and some smaller US cities never had a cocaine epidemic as large as did cities where cocaine use spread before it acquired a bad reputation (e.g., New York, Los Angeles, and Miami). Perhaps the vicarious experience of the problems experienced by those cities served as a "protective factor" for cities that were exposed to cocaine later.

Since the ratio of stage 1 and stage 2 individuals – as well as the number of initiates – varies of time (see Figure 2) it is plausible that optimal demand-side drug control interventions such as treatment and school-based prevention cannot be derived by static optimization models. The application of optimal control theory, however, shows that preventive measures are most efficient when there are relatively few light users typically observed at the onset of an epidemic, while treatment gets more and more important at later stages of the epidemic (see Behrens et al., 2000a). Treatment can even be counterproductive when there are relatively many light users (together with a fast forgetting process). This seemingly perverse result comes from the fact that in this situation the benefits of helping heavy users directly are over-ridden by the cost of not being able to make an example of their suffering. Hence, the efficiency of drug control meas-

[1] We acknowledge that the current cocaine epidemic seems to be over and the next epidemic may be governed by different flow parameters, but we stick with the Everingham and Rydell flow parameters because we do not have data to estimate parameters for any other epidemic and some parameter values are necessary to allow the numerical analysis.

[2] Note that the *LHY* model is consistent with the one presented by Behrens et al. (1999): The parameter-value q changes compared to the *LH* model because the initiation function is augmented with a new state variable.

[3] Note that – at least for cocaine – the relationship between the values of q and δ appears to be approximately linear, i.e. $q=f(\delta)=c\delta$ (c...constant).

ures depends on the time of initial application. Behrens et al. (2000a) found that the optimal budget allocation for prevention and treatment caused interesting dynamics and generated a variety of policy insights such as the benefits of varying the mix of interventions dynamically over the course of a drug epidemic. All these insights allowed to develop a decision-support system.

Additionally, this model underlines Musto's (1987) hypothesis that cycles of drug use may arise when the current generation of youth no longer remembers the adverse experiences of their forebears (see Behrens et al., 2002). We re-discover the old adage that "those who forget the past are condemned to repeat it". Additionally, we verify that it can be desirable to relive past epidemics – at least for myopic decision-makers. Or to put it in simple terms, "*for those who forget the past and over-value the present, it may be optimal to learn the hard way about a particular type of behavior by having their future recreate the past*".

Two-dimensional stability plots can visualize how the qualitative system behavior depends on combinations of two different parameters. One such example might be to look at (1) the discount rate, r, which governs how present- vs. future-oriented the society is and (2) the parameter δ which governs how quickly ever-heavy users are "forgotten". The higher δ is, the sooner past heavy use is forgotten, the less negative is the drug's reputation, and the more likely it will be optimal to recreate the past. As analytically shown in Behrens et al. (2002) this results especially applies for present-oriented societies (large δ, large r).

A parallel scenario for the level of contagiousness (s) and the forgetting rate (δ) is also interesting. (See Figure 3.) As determined in Behrens et al. (2000b) among others cycles may emerge when the parameter s is large, and prevention should help reduce s. Decreasing s, without adapting δ – as represented by the vertical arrow – or vice versa (represented by the horizontal arrow) removes the cycling system behavior and reduces prevalence, both in overall and in equilibrium terms. (Note that the modeled US cocaine epidemic exhibits cycles with a period of approximately 70 years – as observed in the US over the last two-hundred years (Musto, 1987).) That is, moderating the contagious aspect of initiation, e.g., through prevention, reduces the likelihood of cycles and instability.

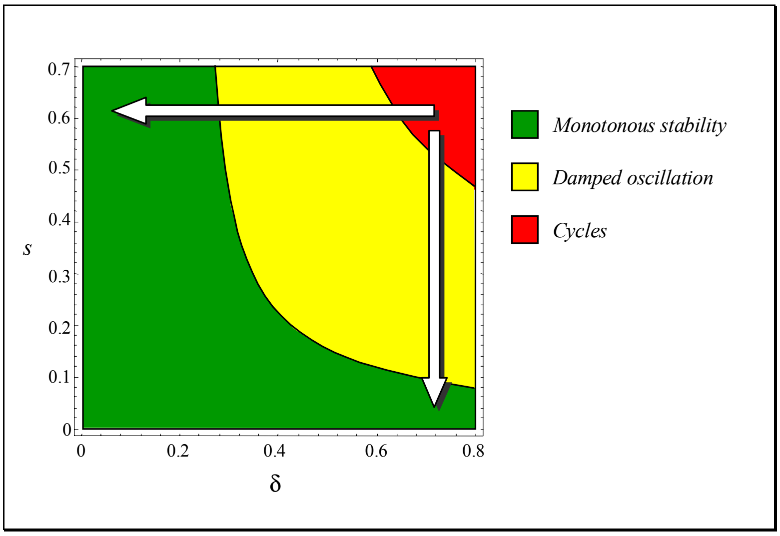

Figure 3: Stability behavior for different combinations of s- and δ-values for a=0.163, b=0.024, g=0.062, τ=50,000 and q=3.443.

Generally, cycling will occur for epidemics with the following characteristics:

- highly contagious (large s), or

- knowledge of the problems of heavy use is discounted (small q) and/or transient (large δ), or
- long average durations of use (large $((1+b/g)/(a+b))$ and addiction (small g), respectively, relative to the persistence and deterrence of memory of drug problems.

Conversely, we can exclude cycling for drugs that have a "well-balanced relationship" between the average duration of use, the addiction career length, and the decay of memories of past problems. I.e. drug epidemics exhibiting this property will decay of their own accord.

The analysis of the model presented here generates a number of additional insights concerning drug control interventions among which are the following: (1) Both the costs of interventions and the social cost associated with drug consumption increase with a delay in the starting year of control giving evidence of the necessity of a "monitoring system"; (2) decision-makers who perceive drug use to be costly for society should favor greater drug control spending per gram consumed and allocate a greater proportion of that spending to prevention; (3) more farsighted decision-makers will rely more heavily on preventive measures (relative to what is done with respect to treatment) to suppress consumption, because this is the most cost-effective way to cope with the drug problem.

From the methodological point of view the analysis of the *LHY* model cries for adding bifurcation analysis to the catalogue of classical OR measures (such as cost-effectiveness analysis).

3. Consequences for decision-making support

Let us now return to the general *LHY*-model. The emergence of stable cycles – as described above – is no pathological exception from the rule. Oscillation emerges for certain parameter sets as described in Behrens et al. (2000b). Combining all insights about the dependence of a behavior's stability and the endemic levels on the choice of the parameter values, we can draw some inferences about the likely effects of several types of interventions.

(1) **Prevention**. For any epidemic, regardless of its special features, it is always useful to reduce the drug's contagiousness (e.g., through peer resistance-based prevention programs), but this is particularly so at the onset of a drug epidemic (when both the number of heavy users and the memory of the years of abuse are small in size). Doing so helps to decrease consumption significantly and to avoid cycles.

(2) **Treatment**. For epidemics of potentially risky behavior with a low forgetting rate it is always useful to help heavy users quit use, e.g., through treatment, though treatment is particularly effective at the end of an epidemic. However, for epidemics with high forgetting rates treatment is not advisable. On the contrary, intensive treatment pursued throughout the epidemic can induce cycling in the system behavior. This certainly sounds cruel and brutal but it means nothing else but that societies immediately eradicating the consequences of particular types of behavior without implementing information campaigns in exchange (= "fighting the symptom not the cause") will hardly have the chance to learn about the negative consequences of the type of behavior in question. Hence, the relationship between the effects of treatment, the efficiency of information and the degree of forgetting has to be investigated in detail before control measures are installed to accomplish an "easy" learning process.

(3) **Relative Timing of Prevention and Treatment**. With this model the extreme finding from Behrens et al. (2000a) that one should never pursue school-based prevention and treatment at the same time is not obtained. However, it remains useful to intensify prevention in the beginning and treatment later in the epidemic.

(4) **Enforcement**. For epidemics with a high probability of users escalating to heavy use, it may be useful to pursue measures that discourage escalation. Using enforcement to keep prices high may be one such intervention. Note that these measures have to be planned in a very careful and well-balanced way, because a low probability of escalation from light to heavy use might cause recurrent behavior (especially when appearing together with a fast forgetting process).

4. Ideas for further research

So far, we mostly limited ourselves to (health) economics applications. Another example in this field concerns the transmission of HIV or hepatitis C through unprotected sexual intercourse. In this case *L* designates the group of those having unprotected sexual intercourse ("risky behavior"). The escalation from *L* to *H* describes the infection as a consequence of this potentially harmful behavior, i.e. H describes the group of HIV or hepatitis C infected persons ("harmful"). The information about such infections represented in the "memory" *Y* reduces the risky behavior with regard to unprotected sexual intercourse (*I*). As opposed to the standard model the escalation rate (*b*) depends in this case on the size of group *H*.

The *LHY* model is so general, however, that it can be applied to all kinds of behavior patterns such as marketing models, for instance. Consider the case of durable consumer goods, such as cars. For this purpose the consumers are roughly divided into satisfied (*L*) and dissatisfied (*H*) customers. While the latter will negatively influence the buying decision of potential new customers (*I*), satisfied customers will positively reinforce their propensity to consume. The traditional marketing instruments like quality, price and advertising of a profit-maximizing company increase *I* and slow down the growth of the number of dissatisfied customers (*bL*).

Innumerable other applications of the *LHY* model are also conceivable in OR but we prefer to leave it up to the readers – not least for space reasons – to discover them by themselves.

Acknowledgements

This research was partly financed by the Austrian Science Foundation under contract No. P14060-SOZ, the National Consortium on Violence Research, and by the US National Science Foundation under Grant No. SBR-9357936. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[Behrens et al., 1999] D.A. Behrens, J.P. Caulkins, G. Tragler, J.L. Haunschmied, and G. Feichtinger: "A Dynamic Model of Drug Initiation: Implications for Treatment and Drug Control". *Mathematical Biosciences*, **159** (1999): 1–20.

[Behrens et al., 2000a] D.A. Behrens, J.P. Caulkins, G. Tragler, and G. Feichtinger: "Optimal Control of Drug Epidemics: Prevent and Treat – But Not At the Same Time?". *Management Science*, **46**(3) (2000): 333–347.

[Behrens et al., 2000b] D.A. Behrens, J.P. Caulkins, G. Tragler, and G. Feichtinger: "Memory, Contagion, and Capture Rates: Characterizing the Types of Addictive Behavior that Are Prone to Repeated Epidemics". *Discussion paper* No. **251**, Vienna University of Technology, Institute of Econometrics, OR and Systems Theory, December 2000.

[Behrens et al., 2002] D.A. Behrens, J.P. Caulkins, G. Tragler, and G. Feichtinger: "Why Present-Oriented Societies Undergo Cycles of Drug Epidemics". *Journal of Economic Dynamics and Control*, **26**(6) (2002): 919–936.

[Caulkins et al., 2000] J.P. Caulkins, M. Dworak, G. Feichtinger, and G. Tragler, G.: "Price-Raising Drug Enforcement and Property Crime: a Dynamic Model". *Journal of Economics*, **71**(3) (2000): 227–253.

[Everingham and Rydell, 1994] S.S. Everingham and C.P. Rydell: "Modeling the Demand for Cocaine". MR-**332**-ONDCP/A/DPRC, RAND: Santa Monica, CA, 1994.

[Knoll and Zuba, 2000] C. Knoll and D. Zuba: "LHY-Model". *Student's project.* Vienna University of Technology, Institute of Econometrics, OR and Systems Theory, 2000.

[Musto, 1987] D.F. Musto: *The American Disease*. Yale University Press: New Haven, CT, 1987.

[NIDA, 1991] National Institute on Drug Abuse: *Overview of the 1991 National Household Survey of Drug Abuse*. White House: Washington, D.C., 1991.

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