Cost-Benefit Analyses of Investments
to Control Illicit Substance Abuse and Addiction

Jonathan P. Caulkins

Carnegie Mellon University Qatar Campus and Heinz School,
5000 Forbes Avenue, Pittsburgh, PA 15213, USA
caulkins@andrew.cmu.edu

Abstract
This paper gives an overview of what is known concerning illicit drug control interventions’ “return on investment” performance from a social planner’s perspective. It is organized by broad type of intervention (supply control, prevention, treatment, harm reduction, and integration across intervention types). The discussion is primarily US-centric, with somewhat greater reliance on international literature vis a vis harm reduction.

Keywords: Drug policy, benefit-cost, cost-effectiveness, dynamic modelling

Acknowledgments
This paper was prepared for the project, "Net Benefits of Public Investment in Individuals: What Do, and Could, We Know?" (David Weimer and Aidan Vining, co-investigators) sponsored by the MacArthur Foundation. Many of the ideas in this essay are due to or were stimulated by discussions with various colleagues, notably Martin Iguchi, Mark Kleiman, Rob MacCoun, Lorraine Mazerolle, Rosalie Pacula, Peter Reuter, and Alison Ritter.
Introduction

The goal of this paper is to give a framework for understanding what is known, what is not known, and what could/should be known about the performance of various illicit drug control interventions from a social planner’s perspective. This “return on taxpayer investment” perspective is sometimes called a “benefit cost” perspective, but in the area of illicit drug abuse much of the work is couched in cost-effectiveness terms. One reason for this is limited ability to monetize the benefits associated with reduced substance use and abuse.

It is common when discussing drug policy to distinguish among four broad types of programs: (1) supply control, (2) prevention, (3) treatment, and (4) harm reduction. There are limitations to this simplistic partition, but for present purposes they are outweighed by its familiarity and transparency, though we add a section (5) discussing interventions that combine or transcend these somewhat artificial boundaries. Finally, there is some (mostly low-quality) literature that purports to quantitatively estimate the consequences of (6) legalization.

It is important to note that the size and quality of the evidence base varies enormously across these six types of interventions (Reuter, 2001). ONDCP (2006) reports the federal government spent $1.04 billion during FY 2005 for research on drug treatment and prevention; so little was spent on research about drug law enforcement that no separate line item was mentioned. That year, the National Institute of Justice (the research arm of the Justice Department) funded $6.4 million in drug-related research. The great majority (69%) of that was for evaluating drug courts and Oxford House treatment programs. Just 24% ($1.6 million) fell under the heading “Drugs and Crime Research”, with half of that share going to Mistral Security, Inc. for the project “Nontoxic Drug Detection and Identification Aerosol Technology”. (The remaining 7% was for evaluating the Bureau of Justice Assistance’s Indian Alcohol and Substance Abuse Demonstration Programs, which include law enforcement, prevention, and treatment components.)

In round numbers, two-thirds of US drug control spending is for supply control, but the federal government invests about 500 times as much studying demand side programs as it does studying supply side programs. This gross imbalance has two implications for benefit-cost studies of drug control. First, relatively modest investments can greatly advance knowledge about the effectiveness of supply control strategies because much low-hanging fruit remains to be harvested. Second, it is harder to meet stringent standards of academic rigor when studying supply control for various reasons, including the paucity of prior research and data upon which to build.

It is also important to note that the “drug system” is complex, so interventions can have many indirect and/or delayed effects that are important, perhaps even larger than the direct effects. E.g., Caulkins et al. (2002) estimate that only one-quarter of a model school-based drug prevention program’s effect on cocaine use comes from preventing program participants from initiating cocaine. Reduced quantities consumed by participants who still use at least some cocaine accounts for another modest wedge, but the majority of the effects may come from spill-over effects on other people through both market and social interactions. Likewise, measures designed to reduce the spread of

---

infectious disease (HIV/AIDS, HCV, TB, etc.) produce spill over effects when preventing one person from becoming infected also averts secondary infections that person would have caused. Naturally quantification of indirect and lagged effects is more tenuous, so the field wrestles with the choice between more rigorous quantification of a subset of effects vs. more speculative quantification of the totality of effects.

A third point pertains to consequences of the scale or scope of interventions. Most treatment interventions are conveniently thought of as operating at the individual or perhaps family level. Hence, it is conceptually if not practically straightforward to imagine collecting data on relevant social welfare outcomes (e.g., employment and earnings, health, etc.) and comparing them to corresponding outcomes for a control group. In contrast, many prevention programs (e.g., media campaigns) operate at the community level, and most supply control interventions seek to alter market conditions. In those cases it is often more natural to evaluate programs in terms of their effect on drug use and then, as a possible final step, to convert reductions in drug use into presumed social benefit using some multiplier (e.g., “an average of $215 in social costs averted per gram of cocaine not consumed”).

Substantial efforts have been invested in extending so-called “cost of illness” (COI) methods to substance abuse and other mental health problems (e.g., Collins & Lapsley, 2002; Harwood et al., 1998). These efforts underpin most such conversion factors, but the COI studies are subject to many criticisms, some correctable and others all but intrinsic to the domain (Moore and Caulkins, forthcoming). Furthermore, crime and violence figure prominently in any policy analysis pertaining to illicit drug control in the US, and there is no consensus on how to put dollar values on crime reductions. Miller et al. (1996) and Rajkumar and French (1997) are among the more widely cited sources for estimates of social costs by crime type, but they are not without limitations.\(^2\) Likewise, it is much easier to calculate the cost to taxpayers per year of imprisonment than it is to estimate the full social cost of imprisonment. Hence, some return on investment studies refrain from using these conversion factors, stopping with a cost-effectiveness result (e.g., kilograms averted per million taxpayer dollars) or apply the conversion factors only as a separate step to facilitate interpretation of the cost-effectiveness results which are viewed as more solid. Not surprisingly, it is evaluations of treatment interventions for which economists are most likely to carry the calculations through to a bottom-line benefit-cost analysis (e.g., French et al., 2002).

The approach adopted here is to emphasize results that appear to be the most important for a policy maker, rather than focusing only on those that are most reliable to a scientific purist or only those conducted within what might formally be construed as a benefit-cost framework.

**Supply Control**

The evidence base is insufficient to support rigorous return on investment analyses of drug supply control programs for at least three reasons. First, it is generally

---

\(^2\) The classic example is property crime (such as theft or burglary). From an academic perspective, economists argue that there is no social cost when stolen goods are merely “transferred” from one citizen (the victim) to another (the offender) without being damaged. Drug policy, however, is the product of a political process beholden to voters who at best roll their eyes at such arguments. Other important examples whose costs are difficult to monetize include homicides, sexual assault, and child abuse.
not possible to conduct experiments, let alone randomized controlled trials (RCTs), particularly of high-level drug enforcement. Second, data systems are weak. Third, supply control is primarily within the bailiwick of law enforcement, and justice agencies do not have the same scientific culture or evidence-based practice norms as does medicine.

Nevertheless, supply control interventions absorb the great bulk of drug control spending in the United States, they set the context (a “prohibition framework”) within which all other interventions operate, and they generate important side effects. They contribute to the very high rates of incarceration among African-American males, with attendant consequences for human capital formation, prison-based spread of infectious disease, disruption of families and parenting, and even in some states disenfranchisement of significant numbers of potential voters (Mauer, 2002). Hence, not attempting some quantification of supply control programs’ performance would be like writing a review for an auto magazine that discussed a car’s sound system but not its engine.

Beyond some early systems dynamics studies (Schlenger, 1973; Levin et al., 1975; Gardiner & Schreckengost, 1987; and Homer, 1993), RAND’s Drug Policy Research Center pioneered attempts to use “risks and prices” (Reuter and Kleiman, 1986) equilibrium market models to estimate how effective various supply control programs could be at reducing US cocaine use (Rydell and Everingham, 1994; Rydell et al., 1996; Caulkins et al., 1997). The basic paradigm was that law enforcement imposes costs on drug suppliers that are passed along to users in the form of higher prices. It is well-established that drug use is inversely related to price (Grossman, 2004) – i.e., the demand curve for drugs slopes downward just as it does for essentially all goods. So such “enforcement taxes” ought to reduce drug use. The RAND studies generally found that while it is cheaper to seize large quantities of cocaine in South America than at the borders and easier to seize large quantities via interdiction than by domestic enforcement within the US, the cost-effectiveness in terms of kilograms of cocaine consumption averted per million program dollars was exactly the opposite, in part because cocaine is so much cheaper upstream in the distribution network that it is easy to replace offshore seizures. Within the US, enforcement against high-level dealers was modeled as having a greater impact on consumption than did enforcement against low level dealers, again in part because low-level dealers are so easy to replace (Kleiman, 1997).

This pessimistic view of interdiction and source country control’s effectiveness was unwelcome news to certain supply control hawks, and a competing study (Crane et al., 1997) soon emerged reaching the opposite conclusions. The studies are not hard to reconcile. One principal difference pertained to assumptions about how cost increases at one market level are passed along to lower market levels. The RAND studies presumed a close to “additive” mechanism whereas the IDA study presumed a “multiplicative” mechanism (so that a 50% increase in the upstream price would lead to a 50% increase in the retail price). See Caulkins (1990, 1994) for clarification of the additive vs. multiplicative models. Second, the IDA study looked at short-term disruptions that were not anticipated by the market; the RAND studies considered effects on the long-run equilibrium of ongoing control efforts to which the markets have had time to adjust.

Nevertheless, a National Research Council review commissioned to reflect on these studies (Manski et al., 1999) dismissed both, and, along with a companion piece
(Manski et al., 2001), argued that the evidence base was at present simply too weak to support scientific evaluation of the benefits and costs of drug law enforcement.

I think the NRC study missed the mark in important respects (Caulkins, 2000), but it nevertheless clearly put a damper on subsequent efforts to quantitatively assess the performance of drug law enforcement from a social welfare perspective.

One of the few studies since then (Kuziemko and Levitt, 2004) addressed what is perhaps the most perplexing observation concerning US drug markets over the last 25 years. The number of people incarcerated for drug-law violations has increased more than ten-fold since the beginning of the Reagan-era drug war, and has now reached 500,000. The vast majority of those imprisoned were involved in drug distribution, not merely drug users (Caulkins and Sevigny, 2005), yet prices of the principal illicit drugs have fallen sharply (Caulkins et al., 2004). Kuziemko and Levitt conclude that prices would have fallen even were there no incarceration not expanded. In particular, they credit the 1985 – 1996 increase in the number of drug-law violators behind bars (which Caulkins and Chandler (2006) estimate to be from 82,000 to 376,000) with raising cocaine prices by 5-15% above what they otherwise would have been. Still, that is far less than “risks and prices” based models would have predicted, so there are still questions about the differences between the empirical and model-based cost-effectiveness estimates (essentially 5 vs. 27.5 kilograms per million taxpayer dollars spent, respectively). Furthermore, uncertainty about the social cost per year of imprisonment, as opposed to its budgetary cost to taxpayers, leaves open the question of what is the overall benefit-cost ratio of this incarceration.

Prevention

Not surprisingly, people who work in the prevention field are excited about prevention programs with “promising” initial evaluations and believe their field is making progress. They speak of having created a new discipline (“Prevention Science”), and stress the gap between what they see as proven “model” programs and what is implemented in practice (e.g., Hawkins et al., 2002). However, a possibly growing number of scholars are increasingly skeptical that prevention has any material effect on drug use.

To over-simplify, one might distinguish three schools of thought concerning prevention. The optimists believe that prevention works, is cost-effective, and that the main challenge is getting the “proven” programs adopted in practice. The skeptics are decidedly negative about prevention and explain the small minority of studies that return statistically significant effects as the inevitable result of running a very large number of evaluations; after all, 5% of evaluations of a placebo will find that it had a statistically significant effect at the $\alpha = 0.05$ level. There is also a middle ground that believes that in certain contexts and for certain behaviors, there are prevention interventions that are cost-effective, but that the average performance is far below what the hype or hopes suggest.

The plurality if not outright majority of prevention evaluations pertain to school-based drug prevention programs, so it is helpful to distinguish school-based programs from all other forms of prevention, which will be addressed more briefly in a separate subsection.

School-Based Drug Prevention
There is an enormous literature evaluating school-based drug prevention programs, but that literature is disappointing in several respects. First, the great majority of studies do not meet standards of quality needed to be trustworthy. Second, many interventions seem to have no effect on drug use, as opposed to knowledge about or attitudes toward drugs. (Notably poor performers are scare tactics, information-based programs, and affective interventions; social skills and “comprehensive” programs seem more promising.) In particular, the most widely implemented prevention program (DARE) has fared quite badly in past evaluation studies, although there is some evidence that DARE Plus is substantially better (Perry et al., 2003), and results from a large evaluation of a revised DARE intervention are due shortly. Third, most studies have short follow up periods (1 – 3 years), and almost none follow subjects beyond the end of high school. Treatment evaluations are also only beginning to follow up clients over longer periods, but a key difference is that treatment clients are already generating substantial social costs. In contrast, the targets of primary prevention for the most part are not yet using heavily, so if prevention does not have long-term effects, then it is unlikely ever to have big effects. Fourth, most longer-term evaluations show decay over time in differences in lifetime prevalence between treatment and control subjects, usually disappearing entirely by the end of high school. There is no consensus, nor indeed any empirical basis, for projecting what such delays in initiation portend for lifetime consumption. Historically, those who start earlier in the absence of an intervention tend to consume more over their lifetime, so it is possible that prevention programs that delay initiation also cause a reduction in lifetime consumption, but that is an assumption/subjective belief, not an empirically established fact.

A recent Cochrane review concluded that skills-based programs (as opposed to those that focus on teaching knowledge about drugs) do reduce drug use (Faggiano et al., 2005). However, with some exceptions (e.g., Plotnick, 1994; Kim et al., 1995; Pentz, 1998), true social welfare studies of prevention are exceedingly rare because the evaluation literature gives information about self-reported prevalence of risk factors (e.g., tobacco or marijuana use) among program participants through the end of high-school, whereas a benefit cost study needs information about actual amount (e.g., kilograms) of use averted of the drugs associated with the greatest social costs throughout the lifetimes of all people affected by the program, including spill-over effects on non-participants. Hence, a social welfare analysis has to bridge quite a few gaps between what is known and what is needed, as suggested by the following table.

Table 1: Gaps Between Traditional Focus of the Prevention Evaluation Literature and the Information Needed to Complete a Benefit-Cost Study

<table>
<thead>
<tr>
<th>Literature tells us</th>
<th>What BC studies need to know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-reported prevalence of</td>
<td>Actual</td>
</tr>
<tr>
<td>marijuana use through 12th grade by program participants</td>
<td>quantity of all drugs (esp hard drugs) over lifetime by all affected by a future</td>
</tr>
</tbody>
</table>

3 The latest study is RWJ funded and headed by Zili Sloboda of the University of Akron who says findings will be reported around the end of calendar year 2006 (personal communication, July 14, 2006).
RCT experiment | scaled up program

Colleagues at RAND’s Drug Policy Research Center and I (Caulkins et al., 1999, 2002) built models to fill in these gaps and thereby support benefit-cost analysis of school-based drug prevention programs. Some pieces were simple (e.g., assuming a 40% hit to effectiveness when moving from supervision by a charismatic program inventor to large-scale bureaucratic implementation). Some were more complicated (e.g., using epidemic models to project a 30% social multiplier “boost” in effectiveness from spill-over effects on friends and associated). The crucial assumption was that 90% of the historical correlation between age of onset and lifetime use carried over to projecting the lifetime consequences of delaying initiation. Collectively, these modeling assumptions generated enormous error bands around projections, to the point that the first book was entitled *An Ounce of Prevention, a Pound of Uncertainty* and the analytic framework was explicitly designed to allow readers to substitute their own beliefs and assumptions and observe the implications on the bottom line.

Nevertheless, to the extent that one is interested in the results of such modeling exercises, the bottom line conclusions were:
1) The point estimate of social benefits from reduced substance use per program participant ($840) was larger than the point estimate of the social costs ($150)
2) That positive BC ratio is quite robust with respect to uncertainty about parameters (though of course does depend on structural assumptions such as the ability to use historical correlations between age of onset and lifetime use)
3) That positive BC ratio comes mostly from prevention being cheap; the absolute effect on lifetime substance use is in the single digit percentage reductions
4) The majority of the social benefits stem from reduced use of alcohol and tobacco
5) Spill over effects on people not in the program may be quite large
6) Social costs were dominated by the opportunity cost of not using classroom time to teach academic subjects.

*Other Prevention Programs*

However unsatisfying is the evidence concerning school-based prevention’s effectiveness and cost-effectiveness, the situation is all the worse for non-school based programs. This is not an area of particular expertise for me, so I will simply quote from the summary of a recent Cochrane review (Gates et al., 2006) of non-school based prevention. That review located only 17 high-quality articles, and they were limited to four specific types of non-school interventions: motivational interviewing or brief intervention, education or skills training, family interventions and multi-component community interventions. The authors’ conclusions were: “There is a lack of evidence of effectiveness of the included interventions. Motivational interviewing and some family interventions may have some benefit. Cost-effectiveness has not yet been addressed in any studies, and further research is needed to determine whether any of these interventions can be recommended.”

Despite the paucity of evidence, three types of interventions not addressed by the Cochrane review merit at least brief mention, the first being media campaigns. The logic behind these campaigns is appealing; harness the same social marketing talents that drive

---

consumer trends, but “in reverse” to discourage consumption of drugs. Furthermore, anti-tobacco campaigns that are in some ways similar have fared well in evaluations (Niederdeppe et al., 2004). The approach has come a long way since the days of the egg-frying PDFA ads, e.g., with messages targeted at parents encouraging them to talk to their children about drugs. In FY 2006, the Office of National Drug Control is investing $99 million in FY2006 in its national media campaign. Evaluating such campaigns is extremely difficult because it is hard to know what “dose” any given individual received, but the evaluations to date of the campaign’s impact on drug use (as opposed to traditional proximate marketing metrics such as recall of having seen the ad) have been negative or inconclusive (Homik et al., 2003), and the OMB notes that it has not been subject to a recent, credible, cost-benefit analysis that shows a net benefit.

The total amount the US as a nation invests in these campaigns is considerably greater than the amount that shows up in the ONDCP budget because the Act creating the federal media campaign mandated corporate matching (donation of advertising time and pro bono offers of services to create the ads), and because of media campaigns run by non-governmental organizations, notably the Partnership For a Drug Free America (to which donated air time is reported to have peaked at $367 million in 1991.

The second non-school prevention modality not covered by the Gates et al. Cochrane review but which merits mention is community level efforts. Perhaps the best known of these is the Robert Wood Johnson Foundation funded “Fighting Back” initiative begun in 1990 in about 15 communities around the nation with annual funding of $2 - $3 million per community. Evaluation results from community-level interventions often are what might be called “process evaluation” or “intermediate outcomes” (pertaining, e.g., to whether the community could organize and what actions were taken). Data on drug use and drug-related problems are harder to gather and analyze. For example, the Fighting Back outcomes evaluation made prominent use of alcohol-related fatalities in the Fatal Accident Reporting System (FARS) database. Community-based interventions are diverse, and it seems likely that one intervention in one community might perform substantially better or worse than another intervention in another community. Yet it is sobering that evaluations of Fighting Back, which might be construed as the “flagship” among community-level interventions, were not terribly positive and, at least in some cases, suggested perverse negative effects (e.g., e.g., Hallfors et al., 2002).

Finally, some observers (e.g., Keith Humphreys of Stanford) note that interventions that are not drug-specific may sometimes have a greater effect on drug use than drug prevention programs per se. Notably, the “good behavior game” is a first-grade intervention designed to improve classroom management, but evaluations have found

---

5 One possible difference is that Florida’s “Truth” campaign depicted the tobacco industry in a particularly unflattering light, stirring resentment against the suppliers, as opposed to merely condemning the product.
6 www.whitehousedrugpolicy.gov/publications/policy/07budget/office_national_drug_control_policy.pdf. Note this is a reduction from the ~$185 million per year level of funding in the early years of the program.
9 “Community” in this context should be broadly construed to include jurisdictions even as large as a state. I am not aware of benefit-cost studies of state-level prevention programs for illicit drugs, but some people credit them with important successes in tobacco and alcohol control.
substantially reductions in hard drug use for program participants relative to controls at junior high school follow up (Furr-Holden et al., 2004).

**Treatment**

Treatment is the best-studied of the various broad strategies for drug control. The conventional mantra is that “treatment works”, and many studies have argued that treatment works not only for the client, in the sense of improving his or her life outcomes, but also in the sense of being a good societal investment (Cartwright, 2000; Harwood et al. 2002). Belenko et al. (2005) recently published a readable and insightful critical review of the economic benefits of drug treatment based on 109 economic evaluations of treatment published between 1990 and 2004. Given the size of this literature and the availability of Belenko et al.’s review on-line, I will not attempt a literature review, but rather will focus on a few key points that help explain and qualify the “treatment works” mantra.

First, it is important to note that, at least for hard drugs, dependent substance abusers create such enormous social costs that even interventions with modest success rates tend to look great from a return on investment perspective. E.g., in round terms, a dependent US cocaine user spends about $10,000 per year on cocaine that costs about $100 per pure gram, implying consumption on the order of 100 pure grams per year. When COI study results are parcelled out by substance (Caulkins et al., 2002), it appears that the social costs are on the order of $215 per pure gram of cocaine. That suggests an annual social cost per dependent cocaine user upwards of $20,000 per year. Annual social costs per opiate user may be even higher given the greater risks of transmitting blood-borne diseases with injected drugs.

Furthermore, residual career lengths are not short, meaning that in the absence of treatment, the ~$20,000 per year social cost would be incurred for many years to come. Studies that follow treatment clients over extended follow-up periods find that of those who do not die, many continue using (Hser et al., 1993, 2001, 2006). Hence, an intervention that managed to persuade a dependent user who would have consumed for another 10 years to cease use permanently would generate a social benefit on the order of $200,000. If the average cost per treatment admission is in the vicinity of $2,000, incremental “success rates” of just a percent or two can be enough to produce a favorable benefit-cost ratio.

Nevertheless, the rosy picture needs to be qualified in certain respects. First, treatment success is not uniform across drugs. We have effective pharmacotherapies and by far the strongest evidence for cost-effectiveness for opiates (such as heroin). There is also reasonably good evidence that treatment “works” with marijuana dependence, although the benefit-cost calculus for cannabis looks different because the social cost per person-year of dependent use is much lower and the untreated duration of dependent use may be shorter.

In contrast, the technology for treating stimulants (cocaine, including crack, and various amphetamines including methamphetamine) is far less developed. Considerable funding and research has been invested in research on pharmacotherapies for cocaine but with no important breakthroughs. So treatment professionals must rely on standby “talk therapies” or innovative non-pharmacological interventions such as “contingency management” (or “coerced abstinence”, which is discussed below).
Although some authors have reported favorable benefit-cost ratios for cocaine treatment (e.g., Flynn et al., 1999), the NRC (Manski et al., 2001) was decisively negative concerning the lack of rigorous evidence concerning cocaine treatment efficacy, holding that there is no solid evidence upon which one could estimate effectiveness let alone cost-effectiveness or a benefit cost ratio. In particular, before and after studies are vulnerable to regression to the mean and other statistical problems, and there is a dearth of solid evidence from randomized controlled trials. Hawken (2006) includes a review of benefit-cost ratio studies of treatment, and almost all rest on before and after designs (including such classics as Gerstein et al., 1994 and Harwood et al., 1998).

One irony is that within the treatment community there is such widespread belief about treatment’s benefits that institutional review boards would generally reject as unethical any study that randomly assigned dependent users to a “no treatment” condition. So it is not clear how evidence would ever be produced in the US that would change the mind of the NRC skeptics.

Instead, many treatment studies look for differences in effectiveness between one program and another. The literature is vast, so no simplistic generalizations can do it justice, but it is not uncommon to find no or at most modest differences in outcomes across programs (e.g., Morral et al., 2006). Skeptics say that is because no program achieves much and apparent benefits are statistical artifacts, e.g., from regression to the mean. The more charitable view is that drug dependence is a general condition with a host of correlated social problems, so the purely medical model of matching the right medical technology to the right specific diagnosis is misguided.

The one possible exception to this general pattern of failure to develop advanced medical “technologies” for treating stimulants abuse is immunotherapies (“vaccines”) against cocaine and methamphetamine, as well as some other substances, including nicotine. These immunotherapies work by having antibodies bind to and inactivate drug molecules. Animal trials and early research with humans is intriguing (Harwood, 2003). Kleiman (2003) offers a preliminary benefit-cost projection based on early data and educated guesses, but updating that work based on the recently completed Phase II clinical trials for cocaine and the rapidly expanding knowledge concerning this technology might be valuable.

The second qualification to the “treatment works” mantra is that many treatment professionals do not like to even think about evaluating programs in terms of proportions of people who become permanently abstinent. They stress that drug dependence is a chronic relapsing condition, akin to diabetes; that one should think of managing that chronic condition, not curing it; and, hence, that most of the benefits accrue while the client is in treatment, not from changed behavior subsequent to and as a result of treatment (McLellan et al., 2005).

In-treatment effects may be enough to cost justify programs. Suppose a program cost $10,000 per year and had 100% relapse rates the moment clients left the program, but that during treatment drug use or more generally drug-related social costs were reduced by 60%. Inasmuch as 60% * $20,000 in social costs per user-year is greater than the $10,000 per year program cost, such a program could be cost-justified. Indeed, a number of benefit-cost studies (e.g., Pollack, 2001) have argued that drug treatment can be a good social investment even if it has no effect on drug use, simply by virtue of reducing social harms associated with that use, notably HIV transmission.
Third, there is no reason to think that the return-on-investment for treatment is the same for all subpopulations. The classic example is the possibility of greater social benefits when treating pregnant women, though given Pollack’s (2000) observations about “crack babies” being primarily a pediatric not an obstetric problem, this logic might extend to all drug-dependent parents of young children, not just women who are currently pregnant. (There are on the order of one million children living with one or more parents who are dependent on illicit drugs.)

Likewise, given that a large proportion of the social costs associated with illicit drug use pertain to drug-related crime, the benefit-cost ratio from treating criminally-involved dependent users may be higher than that for treating drug dependent users more generally. To be more precise, the key distinction is probably not between those who have ever and those who have never committed a crime, since some degree of criminal involvement is so common. Rather, the key distinction is between high-frequency and less than high-frequency offenders. The highly skewed distribution of offense rates (Blumstein et al., 1993) implies that a small subset of criminals are responsible for a disproportion proportion of all offending.

Harm Reduction
Outside the U.S., harm reduction is often seen as a co-equal fourth category of drug control interventions, on a par with enforcement, treatment, and prevention. Two distinctions are important to make with respect to harm reduction. First, harm reduction can refer to an overall philosophy of drug control policy. When viewed in that way, there is little rigorous data concerning effectiveness because it tends to be implemented at the jurisdiction level, often for an entire country, and there is not a relevant comparison group. Cross-national comparisons are instructive at the conceptual and qualitative level (MacCoun and Reuter, 2001), but rarely provide a solid basis for quantitative estimation of return on investment. The potential for quantitative evaluation is greater for specific harm reduction interventions, as opposed to general philosophies of drug control.

Second, when it comes to the harm reduction evaluation literature, one must distinguish between needle and syringe exchange programs (NSP) on the one hand, and everything else on the other hand. There is a large body of literature evaluating NSPs, including return on investment studies. NSPs are not uncontroversial in the United States, with some worrying about “sending the wrong message” or otherwise stimulating drug use, but the predominant consensus in the academic literature is that NSPs are an effective way to reduce the spread of HIV/AIDS and, not surprisingly given the social cost of AIDS, that they are cost-effective. E.g., Ritter and and Cameron’s (2005) recent systematic review of harm reduction interventions finds cost-effectiveness in the range of $4,000 - $35,000 per HIV infection averted. Ritter and Cameron (p.19) also describe the literature as being “predominantly positive” with respect to Hepatitis C (HCV) control, although Pollack (2001) highlights the marked differences between HIV and HCV control. In brief, the HCV virus is so much more robust that HCV prevalence can be very high even in populations following harm reduction practices sufficient to keep HIV prevalence at very low levels.

Ritter and Cameron describe the quantitative, scientific evidence base evaluating all other harm reduction interventions as being much thinner. For example, of the 680 evaluation studies they reviewed, more than half (344) pertained to NSPs. Fewer than
10% pertained to the second most frequently evaluated harm reduction intervention, safe injection facilities (SIFs), which seek to reduce health problems and disorder associated with drug use by allowing drugs obtained elsewhere to be consumed on premises with no risk of arrest and typically with some form of supervision.

I am not personally well-versed with the SIF literature, but based on Ritter and Cameron’s review, they seem to offer limitations and opportunities for cost-effectiveness analysis that are in some ways akin to those associated with drug enforcement. That is, there does not appear to be a rigorous evidentiary base sufficient to produce definitive return on investment studies based on evidence from randomized controlled trials. However, projections of plausible cost-effectiveness ratios based on reasonable assumptions, break even analysis, and/or other modeling analyses might possibly offer policy-relevant insights concerning orders of magnitude effects.

One final observation is that in a second, distinct respect cost-effectiveness analysis of harm reduction may be more akin to prevention. Moore (2005) and Rigter (2006) show that even in countries such as Australia and the Netherlands that pursue harm reduction fairly vigorously, the budgetary costs of harm reduction are minor compared to the costs of other forms of drug control. Hence, it may be that if harm reduction is effective at all (and does not produce perverse results), then it would be cost-effective because the costs are so low that large effects are not needed to pass a cost-effectiveness threshold.

**Integrative Interventions**

Not all interventions or benefit-cost analyses fall neatly into a supply-control, prevention, treatment, or harm reduction bin. For example, I’ve spent a considerable portion of my energy over the last ten years modeling how the cost-effectiveness of different interventions and combinations of interventions might vary over the course of a drug epidemic (Behrens et al., 1999, 2000, 2001; Tragler et al., 2001; Caulkins, 2001, 2004). The typical conclusions are that enforcement is relatively most cost-effective early, in the explosive growth stage of an epidemic, whereas treatment and harm reduction may be more useful later, in the endemic stage.

Drug testing programs are another example, for both employees and students. Testing is common, with about half of full-time employees working for firms that drug test. There is a modestly large literature on drug-testing, which MacCoun (2006) comments on insightfully. There are some quantitative strands, but I do not know of strong ROI studies at the societal or even organizational level, although several colleagues I polled informally reacted positively to the potential value of such a study.

Likewise, Mazerolle et al.’s (2006) systematic review of drug law enforcement found that proactive interventions involving partnerships between police and other entities appear to be more promising than traditional policing interventions. In particular, there are several interventions at the intersection of law enforcement and treatment that have been and/or merit investigation within a benefit cost framework.


Drug/DWI Courts

Perhaps the best known and longest standing partnership between criminal justice and treatment are “drug courts” of which there are now more than 1,200 in the nation. There are many varieties of drug courts, but generally they allow on a case by case basis for non-violent drug offenders to be sentenced to treatment in lieu of the standard prosecution. Drug courts typically involve frequent visits before the court and often urinalysis to monitor the defendant’s progress, with the judge doling out punishments and rewards (e.g., reduced frequency of drug testing) in response to progress or backsliding, with sufficiently poor performance leading to a jail or prison sentence (Shanahan et al., 2004).

Similar to the literature on conventional treatment programs, non-experimental studies that compared program completers to non-completers, to those ineligible for the program, and/or to the clients before the program suggest drug courts offer a very favorable return on investment. However, such non-experimental studies are subject to potentially significant biases that undermine the credibility of their conclusions, as Belenko (2001) in particular has noted. Eibner et al. (in submission) report that experimental and quasi-experimental studies produced mixed results (e.g., Gottfredson et al., 2005 found favorable results, whereas Deschenes et al. 1995 did not).

The popularity of drug courts has spawned similar innovations in other areas. One of particular interest is DWI/DUI courts (Freeman-Wilson and Huddleston, 1999). Total deaths (~17,000) and economic costs (over $100B) associated with alcohol-impaired driving are a bit smaller but still of comparable magnitude to acute deaths and social costs associated with all illicit drug use overall. Substantial progress has been made in reducing DWI, e.g., by raising the drinking age. Progress has slowed, though, and DWI courts offer some hope of further reductions in that substantial social problem. To the best of my knowledge, there has not been a benefit-cost study of DWI courts. Eibner et al. (in submission) may come the closest, examining a DWI court in LA for which outcomes were similar to those achieved with standard judicial processing, but at a lower cost for some offenders (specifically 3rd time offenders).

Given that drunk driving generates social costs roughly comparable to those of all illicit drug use, the paucity of BC literature on DWI courts relative to conventional drug courts suggests that there is relatively greater need and potentially greater benefit from additional economic evaluations of DWI courts.

Prison Diversion or Treatment-in-Lieu of Incarceration

Treatment-in-lieu of incarceration policies, such as California’s Proposition 36 passed in 2000, are like drug courts in that they divert non-violent offenders away from prison while being more intensive than probation without treatment. However, they, or at least California’s Proposition 36, differ in certain respects. Proposition 36 applies across the board to all eligible offenders, rather than being available at judicial discretion on a case by case basis, and under Proposition 36 any discretion concerning treatment for eligible participants (e.g., those without violent priors) is made by medical professionals, not judges. Also, the subsequent monitoring is modest. Indeed, only about 25% of offenders complete their treatment (Hawken 2006). Hawken (2006) has assembled an amazing data set on Proposition 36 and shown persuasively that Proposition 36 saved the

---

state of California money, basically because treatment is much less expensive than prison and that edge persists even with a comprehensive accounting of various indirect costs. However, as of yet there has not been a comparable analysis of Proposition 36 from a social welfare rather than a budgetary perspective. Such an analysis would be timely and valuable.

Coerced Abstinence

Mark Kleiman (1997) has suggested taking the judge and the individual discretion out of drug courts, replacing them with a simple, well-defined and unyielding system of graduated sanctions for failing drug tests. If the offender thinks he or she needs treatment in order to stay clean, treatment can be made available. However, those who can achieve abstinence on their own are welcome to do so. In principle, coerced abstinence can be scaled up to larger sizes than can drug courts because it makes less intensive use of scarce judicial and treatment services. In practice, it is a difficult concept to implement because of lack of coordination between various agencies (courts, probation, prison, etc.) Nevertheless, Kleiman observes that the great majority of the cocaine and heroin consumed in the US is used by probationers and parolees who are nominally under criminal justice supervision. Current case loads make that supervision all but vacuous, but if it were made real, e.g., by frequent testing with certain and swift but not draconian sanctions, there is at least in principle the potential to eliminate the bulk of the problematic drug use in the US. No other intervention can credibly make such a claim.

A notable attempt to implement the coerced abstinence principles occurred in Birmingham Alabama, Jacksonville Florida, and Tacoma Washington. The coerced abstinence model was not fully implemented, e.g., sanctions for failed drug tests were too mild, primarily because of substantial challenges in coordinating across bureaucratic boundaries. Nevertheless, the evaluation (Harrell et al., 2003) found Breaking the Cycle reduced drug use, crime, and family problems, and offered a favorable benefit cost ratio in all three sites.

Apparently, the Hawaii Attorney General’s Office has led a similar project called HOPE (Hawai‘is Opportunities for Probation with Enforcement) that might have been more successful at integrating the various relevant components of the criminal justice system. Colleagues and newspaper reports suggest HOPE has produced promising results (suggestions of declines in drug use among probationers from 30% to 6%), but it has not been formally evaluated, let alone subjected to a true benefit-cost analysis. If so, that would be exciting since, as mentioned, there really is no other drug intervention that offers even the hope of the sorts of reductions in drug use and drug-related costs that coerced abstinence in theory may be able to produce.

Benefit-Cost Studies of Legalization

When a nation bans a psychoactive drug, it creates a black market, and black markets generate a variety of harms, notably crime, corruption, and control costs, but also in some cases increased health harms because the drugs are adulterated or the drug ingestion process is more harmful than it otherwise would be (e.g., more likely to be injected with non-sterile syringes). So prohibition’s benefits of reductions in use that are

---

13 July 24, 2006 Honolulu Advertiser article
caused by higher prices, lower purity, lower availability, social stigma, and/or criminal sanction risk are to some extent offset by a different set of costs or harms. Conversely, legalizing a currently banned substance would avert a great many social costs associated with the black markets but increase use and potentially use-related harms by some unknown extent.

For every important substance except marijuana there is essentially no serious empirical base upon which one could build a benefit-cost analysis of legalization because no modern, developed nation has ever legalized substances such as cocaine or heroin. Several have decriminalized and most have far less stringent forms of prohibition than the US, but none have made production, distribution, possession, and use legal for recreational purposes for at least some important subset of the population). MacCoun and Reuter (2001) have published the definitive study of what can be learned about legalization by studying other times, places, and vices, and they are largely agnostic, cautioning against drawing too firm conclusions. I have made the same point in an entirely different way in a short paper noting that structural assumptions about the nature of the demand curve over price ranges for which we have no empirical data can very dramatically affect projected effects on drug use (Caulkins, 2001). Even though, in my opinion, it is not responsible to pretend one can reliably quantify the benefits and costs of legalizing any of the major drugs, with the possible exception of marijuana, that has not stopped people from trying, particularly economists who are predisposed to believe that free markets reliably maximize social welfare. Examples include Clarke (2003) and Becker et al. (2006).

Benefit-cost analysis of marijuana regime changes, such as decriminalization or depenalization, as opposed to true legalization, are a different matter altogether. They can be on much firmer ground because various jurisdictions in the US and abroad have experimented with alternate regimes. Kleiman (1989) gives an excellent early analysis, and Hall and Pacula (2003) offer the current state of the art in social welfare analysis of marijuana regime change. It is possible to overstate what marijuana decriminalization might produce in terms of reductions in incarceration (Caulkins and Sevigny, 2005) and black market related social costs more generally, inasmuch as most such costs stem from cocaine, heroin, and methamphetamine markets, not marijuana markets. Still, marijuana is by far the most widely used illicit drug in the US and it accounts for the plurality of drug arrests, so it is a drug policy question of first order importance. Furthermore, there is a steady stream of new information whose interpretation is subject to scholarly debate (e.g., recent findings concerning marijuana and schizophrenia, Pollack and Reuter, mimeo). Also, much past empirical work concerning marijuana decriminalization in the US may need to be reexamined given recent work by Pacula and colleagues (2003) who showed that the conventional notion was wrong that eleven US states decriminalized marijuana in the 1970s and that decriminalization meant more or less the same thing in every state is seriously flawed. Hence, the old standard way of modeling decriminalization in statistical analyses (dummy variables set to 0 before and 1 after decriminalization) is inadequate.

Conclusions

There is a modest and growing but immature literature on benefit-cost or more generally return on investment analysis of drug control interventions. It is of very uneven
quality, reflecting differences in quality of the underlying evidentiary base, data, and funding support across different classes of interventions (most notably, enforcement vs. treatment).

Several topics reviewed above were highlighted as promising areas for further benefit-cost analysis. There is at least one other that merits mention even though it was not discussed above because it did not fit neatly into the organizing framework. Abuse of diverted prescription drugs has always been a substantial problem, particularly with regard to emergency department overdose mentions. However, there are signs that the problem is growing, and there is a belief that specific interventions can lead to important reductions in morbidity and/or mortality associated with specific substances (GAO, 2002). Hence, it is plausible that interventions pertaining to abuse of prescription drugs would also be good candidates for further study.

However, it is in general not possible to say “this or that area is the best topic for the next benefit-cost analysis” because what is “best” can depend considerably on whether one values more a rough analysis of an important but understudied intervention or, at the other extreme, yet another strong analysis of a well-studied intervention such as needle and syringe exchange programs or methadone maintenance.

References


Eibner, C.; Morral A; Pacula R.L.; and J. MacDonald. “Is the Drug Court Model Exportable? An Examination of the Cost-Effectiveness of a Los Angeles-Based DUI Court” Journal of Substance Abuse Treatment, Forthcoming.


Godfrey C; Stewart D; Gossop M. Economic analysis of costs and consequences of the treatment of drug misuse: 2-year outcome data from the National Treatment Outcome Research Study (NTORS). *Addiction* 99*(6)*: 697-707, 2004.


Reuter, P., G. Crawford, and J Cave (1988). Sealing the Borders: The Effects of Increased Military Participation on Drug Interdiction, Santa Monica, CA, RAND.

