

Drug Policy Research

Chapter 25 of *Addiction Research Methods*, edited by: Dr. Peter G. Miller, Professor John Strang,
and Professor Peter M. Miller.

Jonathan P. Caulkins

Carnegie Mellon University
Qatar Campus and H. John Heinz III School of Public Policy & Management
5000 Forbes Ave. Pittsburgh, PA 15213

caulkins@cmu.edu

1-412-268-9590 (office)

1-412-268-5338 (FAX)

Rosalie Liccardo Pacula
RAND Drug Policy Research Center
RAND Corporation
1776 Main St., P.O. Box 2138-90401
Santa Monica, CA 90401-2138

Pacula@rand.org

1-310-393-0411 ext 6494 (office)

1-310-393-4818 (FAX)

September, 2008

Acknowledgments

This work was funded in part by the Robert Wood Johnson Foundation, the Qatar Foundation, and the National Institute on Drug Abuse (1R01 DA019993).

Drug Policy Research

Introduction

Drug policy research is the application of policy analysis in the substance abuse domain with a level of rigor that merits publication in academic journals on the grounds that the methods and/or results can provide foundational insights upon which subsequent analyses might draw.

Policy analysis in turn is an interdisciplinary field that strives to objectively and empirically understand the consequences of different public policy interventions, including both retrospective evaluation of past interventions and prospective projections of contemplated interventions.

It is useful to distinguish three types of policy analysis:

- 1) Analysis of net effects on society as a whole (a “social planner’s perspective”),
- 2) Distributive analysis of effects on each significant group of stakeholders, and
- 3) Political analysis of what convergence of forces can push through a piece of legislation or other policy change.

The following example illustrates the differences. Many analysts believe increases in “sin taxes” on tobacco, alcohol, and gasoline would be welfare enhancing. Consumption of these goods generates “externalities”, meaning costs suffered by people other than the producers and consumers. Society would be better off if consumers substituted some consumption into other goods that did not generate such large external costs. A so-called “Pigouvian” tax that “internalizes those externalities” can create the right incentives for people to reduce consumption down to socially optimal levels. That sort of thinking is an example of societal level policy analysis.

However, even if raising excise taxes improved aggregate social welfare, not everyone would benefit; there would be winners and losers and that is where distributive analysis fits in. In the case of alcohol, winners include those who are not killed by drunk drivers or assaulted by inebriated drinkers. Losers include responsible and casual drinkers who must pay more for their alcohol, the brewery/distillery that sells less alcohol due to higher prices, and distillery workers whose jobs come under threat. Distributive analysis projects outcomes not only in aggregate, but also for each significant stakeholder.

From a social planner’s perspective, the cases for increasing tobacco and alcohol taxes are similar. Indeed, the alcohol case may be stronger because tobacco is so effective at killing smokers that they often die before collecting retirement or social security benefits. Furthermore, with the significant exceptions of second-hand smoke and fires, most of the social costs of smoking fall on the smoker, not on other people. Nevertheless, at least in the US it has been easy to raise tobacco excise taxes, while alcohol taxes are rarely increased, so inflation gradually erodes their real value. Why? A political analysis might begin with the observation that in the US, smokers are a relatively poor, uneducated, ostracized, and politically marginalized minority, whereas alcohol consumption is mainstream. Both alcohol and tobacco industry groups try to rally opposition to proposed taxes, but the alcohol lobby has been more successful because most drinkers like their drink, whereas most smokers are intensely ambivalent about the companies whose products have addicted them (Kleiman, 1992). Such discussion of why certain policies do and do not get enacted is the third type of policy analysis.

One contribution of policy analysis can be helping to overcome limitations inherent in less systematic approaches to debating policy. Policy analysis does this in part by providing analytical frameworks that encourage systematic thinking about the issues and the alternative perspectives that might be taken on the issues. Classic examples include MacCoun and Reuter (2001) for illegal drugs and Kleiman (1992) for psychoactive substances more generally. Another key contribution is providing methods for quantitatively comparing the benefits and costs of various policy alternatives, a topic to which we turn next.

Methods for Quantitatively Comparing an Intervention's Benefits and Costs

Over the last forty years there have been increasing demands to view policy interventions as “investments” in the public good and to hold them accountable for yielding favorable returns in the way that private investors expect their financial investments to yield a favorable financial return. The analogy is not perfect. For example, in Australia it is common to recognize a triple bottom line of economic, social, and environmental outcomes. Nevertheless, programs are more likely to garner support if legislators can reassure voters that with respect to some objective, quantifiable metrics, the expenditure of taxpayer dollars is yielding a good “return on investment”.

Many drug policy analyses seek to assess the quality of these investments using economic evaluation methods, employing metrics such as a ‘benefit-cost ratio’ (BC), ‘cost-effectiveness ratio’ (CE), or cost per ‘Quality-Adjusted-Life-Year’ (QALY) saved. There are several such metrics, some with similar sounding names. (See Table 25.1) Thus, it is important to clarify what are their common elements and principal differences.

Cost identification analysis

All of these metrics begin with a ‘cost-identification’ (also referred to simply as cost analysis), which estimates the value of all resources consumed by the program, intervention, or policy (Gold, et al. 1996).¹ Although cost identification might seem like a simple exercise in accounting, important conceptual issues arise that can influence the results. We defer discussion of them until the next section, both because similar issues arise with respect to estimation of benefits and to keep the present focus on contrasting the methods listed in Table 25.1.

Cost offset analysis

Cost offset analysis counts not only costs generated by the program, but also cost savings. A classic example would be providing preventive care that obviates the need for more expensive treatment (e.g., screening for hepatitis, allowing treatment to begin before the disease has progressed). Sometimes the cost offsets (reductions) are bigger than the original program costs, so implementing the program actually saves money. That might be the case, for example, when non-violent offenders are diverted into treatment in lieu of expensive incarceration. In that case, it may not be necessary to take on the more challenging task of quantifying the non-monetary benefits; as long as they are positive, no matter whether they are large or small, a program that saves money is a good investment.

¹ We will henceforth use the term “program” to refer to either a program, policy or intervention that might be considered by the government or treatment/prevention service provider.

Table 25.1: Common Forms of Economic Evaluations Used in Addiction Studies

Name	Acronym	Outcome Measure
Cost Offset	---	Currency
Cost-Effectiveness Analysis	CEA	Any single outcome (e.g. initiation, past-month prevalence, drug-free days, clean urine samples, length of time in treatment)
Cost Utility Analysis / Quality of Life <ul style="list-style-type: none"> • Cost- per quality of life saved • Cost per disability adjusted life year • Health year equivalent • Quality of Life Index 	CUA / QoL <ul style="list-style-type: none"> • QALY • DALY • HYE • QOLI 	Psychometric indicators of quality of life or extension of life with a disability or illness versus shorter life without the illness/disability.
Cost Benefit Analysis	CBA or BC	Currency

While identifying costs and offsets is necessary, program costs are usually not fully offset through future cost savings, so one must also evaluate what these resources contribute in terms of the desired outcome or outcomes. Furthermore, policymakers frequently are forced to choose between alternative programs that accomplish similar goals. That is where the other forms of evaluation come in.

Cost-effectiveness analysis

Cost-effectiveness analysis (CEA) is the most common form of economic evaluation (Drummond et al., 1997; Gold et al., 1996). CEA evaluates costs and outcomes simultaneously by comparing ratios of the cost of a program relative to its effectiveness at producing a specific desirable outcome. For example, if the key outcome were overdoses averted, then

$$\text{CostEffectiveness Ratio} = \frac{\text{Cost with Intervention} - \text{Cost without Intervention}}{\text{ODs averted with Intervention} - \text{ODs averted without Intervention}}$$

Thus the CEA ratio indicates how much one has to “pay” to “purchase” one unit of the benefit (e.g., cost per overdose). Expressed in this format, smaller ratios are better, indicating less expensive ways of purchasing a given benefit. Some literatures prefer the inverse of this ratio so that larger numbers are better (more benefits purchased per million dollars spent). CEA provides a measure of a program’s cost per unit of outcome which can be compared directly to competing projects that produce the same outcome. The biggest limitation of CEA is that many interventions produce multiple benefits in disparate proportions. Hence, one program may have the best CEA ratio when focusing on one outcome, while a different program appears superior when focusing on a different outcome.

Cost-utility analysis

When an intervention generates multiple benefits that are health-related, cost-utility analysis (CUA) can be used to assess effectiveness using a singular outcome (Drummond et

al., 1997; Yin and Forman, 1995). Like CEA, CUA considers the costs and outcomes simultaneously by forming and comparing ratios. The outcome of interest, however, is expressed in terms of expected utility or general well-being with respect to health, not in terms of a single outcome such as overdose. Changes in utility are expressed in terms of Healthy-Year Equivalents (HYE), Disability Adjusted Life Years (DALY), or, most commonly, in Quality Adjusted Life Years (QALYs) (Gold et al., 1996). The QoL approach requires that each health state of interest be assigned a numerical weight that reflects preferences for that health state, ranging from 0 (death) to 1 (optimal health). The effects of an intervention can then be represented as changes in health states where the quality-adjusted score for each health state is multiplied by the expected time in the state and then summed over the expected time of life.

A number of methods have been used to obtain preference weightings for health states. Specialists in this area grapple with issues like the fact that prospectively people think an adverse health outcome (e.g., loss of a limb) will be much worse than do people who have suffered that outcome and had time to adjust to it. Two particular methods that have been widely adopted include The Healthy Days Measures (CDC, 2000) and the EuroQol-5D (EQ-5D) (Rabin and De Charno, 2001; Kind et al 1998). However, neither fully captures the cognitive, behavioral or physical limitations of addiction; nor do they capture all the non-health dimensions of improved well-being (e.g. better family relations, improved employment situation, less criminal involvement). Hence, scientists continue working on new measures of well-being (for example the “quality of life index” or QOLI) that more accurately capture these dimensions for substance abuse disorders (Daley et al., 2005; Jofre-Bonet and Sindelar, 2004; French et al., 2002).

Cost-Benefit analysis (CBA)

There remains another fundamental limitation of CUA when applied to substance abuse, namely that substance abuse creates and drug control interventions avert a wide range of important non-health related costs and benefits. Policy analysts generally incorporate non-health outcomes into their analyses of other societal issues. For example, when analyzing the possibility of damming a river, policy analysts would factor in outcomes such as electricity production, enhanced recreational outcomes, and flood control (including protection of property as well as lives). These are at best awkward to include within a CUA because they are not naturally thought of in terms of health states. Hence, policy analysts working beyond the domain of health interventions generally view cost-benefit analysis (CBA) as the preferred tool. CBA is similar to CUA inasmuch as it too attempts to convert diverse outcomes into a common metric so they can be added up to get a full accounting of benefits. The principal difference is that in CBA that common metric is currency (dollars in the US, euros in the Eurozone, etc.).

Just as with CUA, there is an entire subliteration devoted to the question of how one assigns reasonable values to various outcomes. This is easy if the outcomes are priced directly in a market, and in modern industrial societies a remarkable range of things are. Nevertheless, a number of key drug-related outcomes are not traded in markets or are traded in highly imperfect markets (e.g. safety, family bonds). Social policy researchers in the addiction field spend a lot of time debating how to deal with these complications. The box on the facing page briefly summarizes the debate surrounding how best to value some of these outcomes.

BOX: Outcomes Whose Valuation Poses Special Challenges

Property crime: What is the societal cost of a burglar stealing something? The old answer is zero on the grounds that theft merely transfers wealth from one person to another, with no net loss to society. Sometimes allowance is made for the possibility that someone who actually bought an object might value it more, on average, than someone who got it for free (stole it). That still skirts the real issue, namely that the public, and policy makers accountable to the public, vehemently object to the idea that there is no social cost to property crime. The more modern approach is to focus on the “willingness to pay” (WTP) to avoid such criminal victimization. That is, if an average person were willing to pay \$1,000 to avoid being burglarized (as evidenced, e.g., by analysis of investments in household alarm systems), then preventing a burglary should be valued at \$1,000.²

Violent crime: A similar conundrum surrounds violent crime. The old approach counted primarily victims’ hospital bills for treating injuries and days of work lost while recuperating, but as with property crime, this led to valuations that struck the public as absurdly low. Estimates based on jury awards and willingness-to-pay principles can incorporate the “intangible” costs of pain and suffering as well as the much smaller “tangible” costs associated with receiving medical care.³

Premature death: How does one value a human life, such as one saved by a drug policy intervention? An often contentious measure employed in many studies is the human capital approach, where the value of a human life is calculated as the present discounted value of expected future earnings (Rice, 1967). A central criticism of this approach is that individuals who choose not to work (e.g. stay at home parents) or those who earn lower wages (e.g. workers in developing countries) are presumed to have less worth than those who earn more, and this seems morally wrong. The willingness to pay approach has been adopted by economists to get around this issue. It calculates the statistical value of life (SVL) based on market purchases people make to reduce their risk of death, or premiums they must be paid to accept additional risk of mortality.⁴ Examples of such market transactions include the purchase of safety devices (cars with airbags, home security systems) and wage premiums for working in risky jobs (e.g. mining or construction). These measures, however, remain contentious as they too can be influenced by the wealth and income of the population in which they are assessed.

Cost of addiction itself: Conventional economic reasoning holds that when people buy something there is a “consumer surplus”, meaning, the monetary valuation of the satisfaction derived from consuming the product must exceed its purchase price, otherwise the consumer would not have freely chosen to pay for it. Non-addicted users presumably are choosing to consume drugs of their own volition, but most of the expensive drugs (heroin, cocaine, and amphetamines) are used by people who are dependent. Are those dependent users more like a typical consumer or more like a prisoner, compelled by the addiction to consume something from which they do not derive much satisfaction? To date, many economic

² Miller et al. (1996) is a classic reference to this more modern approach.

³ Again, Miller et al. (1996) is a classic reference.

⁴ See Viscusi and Aldy, 2003 for a review.

analyses still assume that even dependent users are like conventional consumers and so recognize no cost of addiction per se. Likewise, very few consider the psychological pain and suffering addiction causes the family and friends of those who are addicted because the unit of analysis in many economic studies is the household, not the individual, so the entire household's consumption decisions are assumed to be welfare-enhancing for the household as a unit. Some suggest that when the purpose of the economic analysis is to compare drug interventions to interventions in other policy domains, all analyses must be done using the traditional, common methods. However, when the analysis is only meant to inform choices within the domain of drug policy, then the standard methods should be adapted to domain-specific idiosyncrasies, such as a need to discard standard assumptions about consumption always enhancing the consumer's welfare.

Issues that Arise In Quantifying An Intervention's Benefits and Costs

Box 1 discusses challenges in valuing outcomes that are related to addiction. Distinct from those difficulties is a set of conceptual issues that complicates analyses regardless of one's position on the value of a life or the cost of crime. These include issues related to the perspective from which costs are evaluated; the costs and benefits considered; and the time horizon over which they are evaluated (See Table 25.2).

Perspective

Perspective refers to whose shoes the analyst steps into when deciding which outcomes to include. Common perspectives for economic analysis include the payer, the provider, and society as a whole. When private insurance pays for treatment provided by a hospital or non-profit treatment agency, these three perspectives are all distinct. When the government provides services in its own facilities (e.g., treatment provided through the US Veterans Administration or the British National Health Service), the three perspectives are one and the same. However, even when the public sector is both payer and provider, perspective can still matter if costs borne by one department or ministry produce benefits or cost savings for another department or ministry. For example, when treatment reduces criminality and, hence, associated law enforcement costs, the Ministry of Justice rarely expresses its appreciation by writing a check to the Ministry of Health, so the payer/provider perspective (the Ministry of Health) differs from the social welfare perspective (which would encompass both ministries).

In addiction services, costs have traditionally been evaluated from the payer perspective. Because so many services and programs are paid for by the public sector and because they often have impacts beyond the payer and provider, a U.S. Public Health Service panel now advocates that economic evaluations be conducted from the societal perspective, rather than just from a payer or provider perspective. While the societal perspective has been widely adopted for cost-of-illness studies (ONDCP, 2004; Single et al., 2003; Collins and Lapsley, 2002), some doubt this is truly desirable when considering programs that will only be adopted if they are financially viable for the agency delivering them (Humphreys et al., 2008). Unfortunately, agencies are not frequently evaluated in terms of the savings they generate for other agencies; they are usually evaluated only in terms of costs and performance within their own silo. So, the choice in perspective may depend somewhat on the goal of the evaluation. For an academic audience, the societal perspective may be

preferred, but if the goal is to inform real policy makers, a more local perspective (e.g., that of a single agency or provider) may be more influential.

Economic versus Accounting Costs

Another issue to consider is whether resources should be evaluated in terms of their cost in that use or in terms of their full economic value. Payers and providers have traditionally focused on the accounting or monetary cost of the goods and services they purchase. However, sometimes the price paid for a resource may be more or less than its true market value. A good example would be supplies or volunteer hours donated to a non-profit provider. Although the provider’s cash flow accounting statements will contain no entry for donated items, economists prefer to value these resources in terms of their full value (referred to as “opportunity cost”), as these resources could have been used for other things. The argument is that resources are inherently limited in society, and when they are used for one purpose, they can no longer be used for another purpose. For example, a volunteer who donates time assisting patients in a neighborhood treatment clinic might have a paid job to do similar activities in a hospital. The wage she could have earned if she did a similar job represents the opportunity cost of her time or true economic value of the labor. Similarly, when a teacher uses class time to teach students drug prevention, this represents a real opportunity cost since that time could have been used to teach traditional academic subjects. So the cost of the drug prevention program, to economists, includes the cost of time taken away from other subjects in addition to the monetary cost of the program materials and teacher training.

Table 25.2: Some Issues to Consider When Identifying Costs and Benefits

Perspective	<ul style="list-style-type: none"> • Payer • Provider • Government (Tax Payer) • Society
Costs and Benefits Included	<ul style="list-style-type: none"> • Program costs <ul style="list-style-type: none"> - Accounting cost - Economic (opportunity) cost • Client costs and cost savings • Intangible costs and benefits
Time Period	<ul style="list-style-type: none"> • Evaluation period • Discount rate

An economic evaluation may also consider costs incurred by the individual receiving the program (i.e. the client’s costs). Individuals who have to miss work or need to employ daycare to watch children in order to attend treatment incur real costs to participate in the program above and beyond any monetary co-payments. If the goal is to consider costs from a societal perspective, then these additional costs should be considered. One challenge when moving beyond accounting costs is that it can be like opening a Pandora’s Box of indirect effects that are hard to quantify. Suppose someone attending a thirty-day inpatient program lost their job because they did not have enough accrued vacation and did not want to or could not claim short-term disability. Or suppose that inpatient treatment created strain on a family due to 30 days of separation. In theory, one should include such costs, particularly if the

corresponding long-term effects of improved employment and family functioning are included on the benefit side. Analysts with an agenda can (perhaps unconsciously) distort results by judging that the evidence base concerning some costs is insufficient to support quantification while patching together estimates for other, favorable, effects for which the evidence base is no better. Likewise, bias can enter one step earlier if the program evaluation works harder to document anticipated program benefits than unintended adverse consequences.

Time Horizon

A third issue is the time horizon over which a program's cost and benefits are to be evaluated. Although formal drug treatment (and program costs) might end after a 30 day period, the benefits of that drug treatment can continue to accrue for months and years after the treatment episode has been completed (Zarkin et al., 2005). If a drug treatment program is assessed only in terms of its costs and benefits for the period in which treatment took place, the costs could easily exceed the benefits. However, simply extending the evaluation period by six or nine months can dramatically change the net result. Likewise, targeted law enforcement initiatives like reductions in the availability of precursor chemical laws might have large effects if evaluated within a relatively short time horizon (e.g. months within their adoption), but have no lasting effect when evaluated over a longer time horizon (e.g. two years post). Thus, consideration of the time horizon is important when measuring effectiveness and will likely be partially dependent on the policy being evaluated.

When evaluations are done over periods longer than one year, it is important to make sure that both outcomes (benefits) and costs are discounted to reflect their net present value. Choice of discount factor can have important implications if the horizon over which things are to be evaluated is long, as high discount rates will reduce the benefit (loss) associated with outcomes far off in the future while low discount rates will increase their values.

Methods for Estimating an Intervention's Effects

The previous sections explained how policy analysts aggregate quantitative projections of an intervention's beneficial and costly outcomes into a summary performance measure. The next question is, how does one project program outcomes in the first place? The general answer is, "By any means necessary" or, perhaps more precisely, by any accepted, replicable, objective method that is suitable to the context and evidence. In other words, policy analysis is happy to borrow tools from statistics, econometrics, epidemiology, operations research, or any other science.

This ecumenicalism with respect to methods is admirable; it lets policy analysts grapple with a wide range of issues. However, it is also a source of misunderstanding and sometimes even mistrust from scientists who work within a discipline where one method dominates. For example, some disciplines view randomized controlled trials (RCTs) as not just the gold standard but almost as the *only* reputable form of inquiry. However, it would not make much sense to use randomized controlled trials as the basis for estimating the benefits of space exploration.

Nevertheless, there are good reasons why RCT is the gold standard, and other methods should be employed only when an RCT is not feasible. For treatment and other interventions aimed at dependent users, policy analysis usually builds directly on a clinical

RCT. For example, a group of imprisoned drug users are randomly assigned to standard treatment in prison, treatment plus half-way house support after prison, or a no-treatment control condition. Study subjects are followed up at various times, perhaps 6, 12, and 18 months post-release. Health and criminal justice outcomes might be tracked through administrative data and/or by self-report. The effect of the intervention is judged to be the difference in the average number of events of a particular type experienced or reported by the treatment group relative to the control group (sometimes only recording such differences when the difference is statistically significant). There are limitations to how broadly RCTs can be applied in social research, however. Some interventions, including media-based prevention campaigns and community coalition building, operate at the community level, not the individual level. It is hard to randomly assign communities to one treatment condition or another, and sample sizes when this is done are generally smaller. The problem becomes even more acute when the intervention's impact would be felt nationally, as with changes in national sentencing policy, border control, or interventions undertaken in source countries (e.g., Colombia or Afghanistan) with the goal of constraining supply in final market countries.

RCT designs also do not always measure long-term and spill-over effects. School-based drug prevention programs typically address youth 10-14 years old, yet most of the social costs of drug use stem from use by people in their 30's. Twenty-year follow up data collection is both expensive and rare. Even when possible, it inevitably produces dated information and usually suffers sample selection bias (because people who stay in a study for 20 years are different than those who drop out). The challenge with spill-over effects is that they can be diffuse and manifest in the behavior of people not in the program. For example, cracking down on drug dealing in a particular neighborhood might yield tangible, readily apparent benefits for that neighborhood not so much by reducing drug use as by displacing it. If the displaced activity were spread over ten other neighborhoods (or into social network based rather than place based selling) the displaced activity might be hard to detect, leading to an over-estimate of the net benefits of intervening. Similarly, if harm reduction programs directed at dependent adult users had an adverse effect on initiation by teens, it would be very hard to detect that with typical RCT designs that focus on outcomes experienced by program participants.

Quasi-Experimental Methods

Although randomized experiments remain the gold standard, observational studies using nonrandom population data are often necessary due to practical considerations. In observational studies, systematic differences can exist between the "treated" and "untreated" groups with respect to both observable and unobservable characteristics. Hence, direct comparisons of the observed outcomes from the two groups may not be appropriate and can lead to biased or misleading conclusions. Fortunately, quasi-experimental methods have emerged from a variety of disciplines to assist with the identification of program or policy effects when randomization is not possible. The methods most useful for addiction research include propensity score methods, instrumental variable techniques, difference-in-difference estimation, and regression discontinuity approaches.

Table 25.3: Statistical Tools for Evaluating Outcomes in Policy Analyses

Tools for evaluating program effects:

1. Random assignment study design
2. Quasi-Experimental Methods
 - propensity score (PS) methods
 - Instrumental variables (IV) techniques
 - Difference-in-difference (DND) techniques
 - Regression discontinuity approach
3. Modeling approaches
 - Microsimulation
 - Agent based modeling
 -

Propensity Score Methods. Propensity score methods (e.g. propensity score weighting, nearest neighbor matching) are an increasingly common approach for handling biases associated with nonrandom assignment to programs, policies, or treatment conditions. The methods involve adjusting the comparison group members in terms of their observational characteristics so that they become observationally equivalent to the treated population before the policy (program, treatment) takes place. Assuming no unobserved factor differs systematically between these two groups once this adjustment takes place, analysis of these adjusted data provides a good estimate of the effect of the treatment on the treated population. A variety of methods for adjusting the data have emerged, from “matching” algorithms making use of a small subset of the comparison group to “reweighting” methods that make full use of the comparison group by adjusting the distribution of multiple observed characteristics across the two groups so they are nearly identical.

Instrumental Variables (IV) approaches. IV methods are frequently used for inferring causal associations from observational data, particularly when the explanatory variable (right hand side policy variable, in this case) is known to be correlated with the error term (i.e. is “endogenous”). The endogeneity results from common unaccounted for third factors that influence both participation in the program / policy and the error term (hence, the dependent variable). Standard regression estimation of a model with endogenous variables leads to biased estimation. The advantage of IV methods is that they generate statistically consistent estimates of beta coefficients, enabling the researcher to answer the same question as regression methods (i.e. “what’s the average effect of the program?”). IV techniques depend critically on the ability to identify an instrument that: (1) is highly correlated with the endogenous explanatory variable, conditional on the other variables in the model, and (2) is correlated with the error term itself or directly influences the outcome (dependent) variable being modeled (except through the explanatory variable being instrumented). IV estimates are usually obtained through two step modeling. In the first step the endogenous variable (e.g. program participation) is estimated as a function of all the other variables and the instrument(s) used for identification. In the second stage, the original regression of interest is

estimated but instead of putting actual program participation in the regression, the analyst replaces that variable with the predicted value obtained from the first stage regression. Occasionally reduced form methods are used where the instrument replaces the policy or program variable in the second regression. This allows the researcher to determine if a causal association exists, although they can not then determine the true magnitude of the association.

Difference-in-Differences (DD) Methods. DD techniques attempt to ascertain the average treatment effect of a program or policy by looking at how changes in the outcome of interest differ between the treatment and control groups. Because treatment and control groups could differ in both observable and unobservable ways, it is critical to have data on the outcome variable of interest for at least three different time periods: (1) some time before the policy/program took place, (2) right before the policy/program took place, and (3) some time after the policy/program. With information from these three points one can look at the average change in the outcome of interest for both groups before the policy or program (to see if they had similar rates of change even before the policy took place) and after (to see if the program or policy influenced the rate of change in the outcome variable after it took place). By examining changes in the control group as well as the treatment group, analysts can isolate the true program effect by taking out the average change that would have taken place (based on the controls) even if the program or policy did not take place. Simple examination of mean differences and sophisticated regression methods have been used to construct DD estimates, but the principal idea is the same: to understand if the program or policy influenced the rate of change in the outcome variable for the treatment group beyond what it would have otherwise had the program never taken place (known from the pre-program data and the overall changes observed for the control group).

Regression Discontinuity Approaches. Like DD methods, regression discontinuity methods involve a pretest- posttest group comparison. Groups, however, are assigned not by the program or policy but rather by the analyst based on an observed covariate that is expected to be impacted by the program or policy and that has values on either side of a critical fixed threshold (e.g. the poverty line, 0.08 blood alcohol content level, or possession of 1 ounce of marijuana). Observations in the data are assigned to groups based on the cutoff score of the measured covariate. Observations below the cutoff being assigned to one group, and observations above the cutoff are assigned to the other group. The effect of the program or policy is estimated by the disconnect in the regression lines (slopes) or functions (for higher order polynomials) obtained from separate models estimated on each of these groups below and above the threshold value. Unlike DD, regression discontinuity does not provide an estimate of the average effect but rather identifies whether a difference in average behavior exists at the critical threshold value. Thus, it can identify that a program or policy had an effect, but it can not provide a good estimate of the average effect of the program or policy on the populations as a whole.

Modeling Methods

Quasi-experimental methods can be useful for evaluating past interventions when suitable data are available. However, sometimes policy makers want to project the results of trying something new or replicating something for which suitable data have not been collected in

the past. In such cases, modeling may still be possible if researchers have a good understanding of the structure of the system and the cause and effect rules that it follows. Modeling is rather common in engineering. Space scientists do not estimate how much fuel a rocket will need by collecting data on how much fuel rockets used in the past and running a regression; they build a model of the rocket's dynamics from first principles, such as Newton's laws of motion.

Individual people are rarely as predictable as rockets, but sometimes they behave predictably in aggregate. For example, queueing theory provides a set of models that might do a very good job of predicting how expanding the number of treatment slots would reduce waiting times for people seeking treatment (Kaplan and Johri, 2000). Likewise, product diffusion models from marketing may help predict the spread of a drug that has recently become available within a particular population (Caulkins, forthcoming), and models of infectious diseases may help predict how certain drug control interventions will affect the spread of hepatitis or HIV (Kaplan, 1985; Pollack, 2001, 2002). So-called "risks & prices" models of drug markets try to estimate how changes in law enforcement might affect retail drug prices in equilibrium (Reuter and Kleiman, 1986) and, when coupled with estimates of the "elasticity" (responsiveness) of demand to price changes, how prices changes could affect drug initiation and use (e.g., Caulkins et al., 1997).

Short descriptions of two of the many modeling methods follow.

Microsimulation models. A microsimulation is a computer program that projects the behavior of an individual entity (typically a person, but could be a household or business) within a "synthetic" environment, conditional on that entities characteristics (e.g., gender, age, income, criminal proclivity, etc.). The simulation is then run for a large number of heterogeneous agents whose characteristics match some population of interest (e.g., a birth cohort or the population of a city). The goal is to weight and aggregate up the individual agents' outcomes to understand results for the population as a whole (in terms of average behavior or distribution of outcomes). Running the entire process twice, with different environmental conditions perhaps representing different policy regimes, produces an estimate of how that change would affect population level outcomes. Static simulations explore a fixed population that does not change. Dynamic models can "age" the population for many years (decades) into the future, applying appropriate rules for demographic and economic processes, such as death rates, high school completion rates, retirement rates, and so on. Thus, dynamic models allow one to assess both the immediate and longer term effects of a policy or program on the population as it changes over time.

Agent based models. Agent based models (ABM) are similar to microsimulation models, except that all of the individuals (called "agents") move through the simulation at the same time, and interact with each other, as they follow their behavioral rules (Agar and Wilson, 2002; Perez et al., 2006). For example, in ABM models of traffic flow, the agents are drivers, a general rule common to all drivers might be "drive along the quickest route to work", and individual characteristics might include where the individual agent lives and works. Because the agents interact, and in many models can even learn from these interactions over time, the population level outcomes are not just the summation of the projections for individuals. The agents' behavioral rules can be quite complex, combining elements of game theory, social contagion, and complex systems. Likewise, in data rich

environments (e.g., traffic modeling), ABM models can give quite detailed projections. More often though the goal is to discover “emergent” system behaviors that manifest even when the agents are modeled in a very stylized fashion and follow only very basic rules. Thus, agent based models allow us to examine dynamic systems with complex interactions in a way that other models cannot (due to the inability to track all of the possible interactions).

Summary

Policy analysis strives to understand (often quantitatively) the effects of implementing policies and programs by projecting their outcomes, both intended and unintended, in aggregate and from the perspective of each relevant stakeholder. There are a variety of particular techniques associated with policy analysis, including cost-effectiveness analysis, cost-utility analysis, and cost-benefit analysis. All strive to provide summary metrics that concisely capture and balance the good and the bad consequences of a past or planned action. Hence, policy analysis builds a bridge between scientific understanding of a policy domain and evidence-based decision making concerning how best to intervene in that domain.

Policy analysis in the addictions field presents a number of special challenges that make this type of research exciting, and its findings sometimes controversial. Addiction-related issues span many levels of aggregation, from the molecular, to the individual, agency, community, national, and even international level. Some involve relatively simple logic models; others involve complicated endogeneity (e.g., market behavior), lags, and/or nonlinear feedback. Hence, the best policy analyses are often conducted by interdisciplinary teams that can usefully borrow and adapt methods from many different disciplines.

Further Reading

Bardach (2000) offers the classic, readable introductory text on policy analysis. Stokey and Zeckhauser (1978), Weimer and Vining (2005), and Kleiman and Teles (2006) are also highly recommended. Policy analysis is sometimes best learned by example. The exercises below give the reader a chance to try doing some policy analysis for well structured questions, and we close with references to some organizations and web sites that are good sources of policy research that can serve as models for further work.

Academic Journals and Conferences

International Society for the Study of Drug Policy (www.issdp.org)

International Journal of Drug Policy (through Elsevier, www.elsevier.com/)

Independent Research Organizations

RAND’s Drug Policy Research Organization (www.rand.org/multi/dprc)

National Drug and Alcohol Research Center (www.ndarc.med.unsw.edu.au/)

Including the Drug Policy Modeling Program (www.dpmp.unsw.edu.au/)

International Policy Agencies

United Nations Office on Drugs & Crime (www.unodc.org/unodc/)

European Monitoring Centre on Drugs & Drug Addiction (www.emcdda.europa.eu/)

National Policy Agencies

U.K. Home Office (drugs.homeoffice.gov.uk/)

U.S. Office of National Drug Control Policy (www.whitehousedrugpolicy.gov/)

On-Line Libraries, Bibliographies, and Other Sources

Schaffer Library of Drug Policy (www.druglibrary.org/toc.htm)

Substance Abuse and Mental Health Data Archive (SAMHDA)

(www.icpsr.umich.edu/SAMHDA/)

Substance Abuse and Mental Health Services Administration's (SAMHSA's)

National Clearinghouse for Alcohol and Drug Information (NCADI)

(<http://ncadi.samhsa.gov/>)

Exercises:

1) (*Stakeholder identification*)

Give one example each of a policy that would be good for society that has not been implemented because it lacks political support and a policy that is bad for society overall but which has nonetheless been implemented because a coalition of stakeholders pushed for it. In each case, list the stakeholders who are or would be helped or hurt by that policy.

2) (*Discounting future outcomes*)

a) Imagine two hypothetical policies (A and B) that are mutually exclusive. Policy A will save 100 lives over the next 24 months, and Policy B will save 200 lives in 10 years. Which would you prefer?

b) Now consider two policies that would have no effect on health outcomes, but the first (Policy C) will save taxpayers one hundred million dollars (or pounds, or euros) over the next 24 months, and the other (Policy D) will save \$200 million in ten years. Which would you prefer?

c) For the money saving policies (C and D), find your indifference point. That is, if you preferred C (saving \$100M soon), how large would the savings in 10 years under Policy D be in order for you to be indifferent between choosing Policies C or D? Similarly, if you initially preferred D, how much could the \$200M figure be reduced before you would become indifferent between choosing Policies C or D?

d) Repeat exercise c but for Policies A and B, for which the outcomes are lives saved.

e) Do you get the same numerical value of the indifference point in parts (c) and (d)? Many people do not, but Keeler and Cretin (1983) prove that policy analysis must make intertemporal trade-offs consistently for monetary and non-monetary outcomes to avoid logical traps.

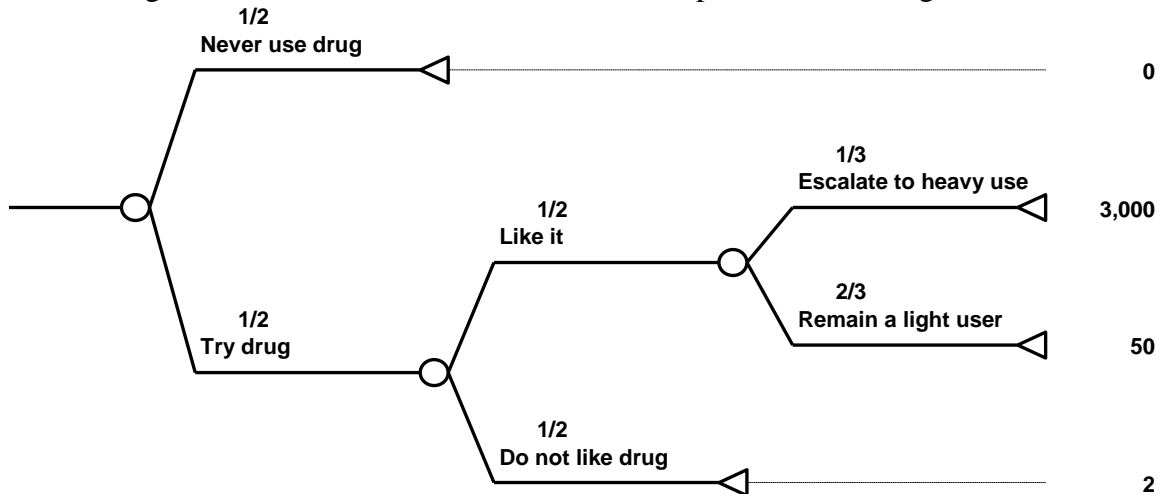
3) (*Event trees*)

There is enormous heterogeneity in drug consumption “careers”, with most people using drugs only occasionally and a small subset progressing to prolonged, frequent use. A full model of this would capture dynamics over time and the use of different types of substances, but many of the policy implications of heterogeneity can be elucidated even by a simple event tree such as the one in Figure 25.1. Circles in event trees represent random events. For instance the first circle in Figure 25.1 represents a situation in which half of all people in the

population of interest (e.g., a birth cohort) try drugs (lifetime prevalence is 50%). Probabilities on subsequent branches are conditional, so the probability of reaching any given “leaf” (the tip of a set of branches on the far right) can be obtained by multiplying the probabilities along the path leading to that leaf. For example, in Figure 25.1, $(1/2) * (1/2) * (1/3) = 1/12$ of the population will escalate and use an average of 300 days per year for 10 years. Leaves are customarily annotated with one or more outcomes. In this case the outcome is the total number of times drugs are used over the lifetime.

- How many times does the average person in this population use drugs?
- Among those who ever use drugs, what is the average number of times drugs are used?
- Show that 96.6% of drug use is attributable to the 1/6 of all users who escalate to heavy use. That is, heterogeneity in this model is even more extreme than the classic “Pareto Law” that says that often 80% of an activity is attributable to the 20% of people who participate most actively.
- How many times more use would be prevented by convincing a light user not to escalate (secondary prevention) vs. preventing a typical person from trying drugs the first time (primary prevention)?

Figure 25.1: Event Tree for Exercise #3: Simple Model of Drug Use Careers



4) (*Breakeven analysis*)

Models of drug use careers become even more useful for policy analysis if they can be augmented by estimates of the social cost or harm associated with each state, transition, or event over the course of that career. To illustrate this, suppose that in the previous problem’s model, every day of heavy drug use generated \$100 in social costs, and costs associated with other drug use are negligible. This implies that the social cost per year of heavy use is \$30,000, which is not an unreasonable figure for dependent use of stimulants, particularly by people who are criminally involved.

- Suppose a universal prevention program cost \$100 per student. What proportion of participants would the program have to persuade not to use drugs in order for the program’s

benefits to equal its costs? (For now, ignore discounting and assume that those whose initiation is prevented are typical of all users in the population.)

b) How would your answer in part a change if prevention only worked on people who would not have escalated to heavy use anyhow?

c) How would your answer in part a change if the program is administered to 13 year olds, the average year of heavy use occurs at age 28, and future events are discounted at a compound annual rate of 5% per year?

d) Now consider a treatment intervention that costs \$2,000 per person admitted. If, on average, people enroll in the program half-way through their career of heavy drug use and the only program outcome were permanent abstinence for a subset of those treated, how high would that treatment effectiveness rate have to be in order for the program to have a favorable benefit-cost ratio?

e) Today treatment programs rarely think in terms of achieving permanent abstinence. Instead, consider a treatment program that costs \$1,000 per month and whose only effect is to reduce use during treatment (no long-term effects). By what proportion would such a program have to reduce use in order to have a favorable benefit-cost ratio?

f) Some interventions seek to reduce harm, not use. If such a program cost \$1,000 per month, by what proportion would it have to reduce harm in order to have a favorable benefit-cost ratio?

References:

Agar, M. H., & Wilson, D. (2002). Drugmart: Heroin epidemics as complex adaptive systems. *Complexity*, 7(5), 44-52.

Bardach, Eugene (2000) *A Practical Guide for Policy Analysis: The Eightfold Path to More Effective Problem Solving*. Chatham House Publishers, New York

Caulkins, Jonathan P. Forthcoming. "Implications of Inertia for Assessing Drug Control Policy." *Contemporary Drug Problems*.

Caulkins, J. P., Paddock, S., Pacula, R., & Chiesa, J. (2002). *School-Based Drug Prevention: What Kind of Drug Use Does it Prevent?* (RAND MR-1459-RWJ). Santa Monica, CA: RAND.

Caulkins, J.P., Rydell, C.P., Schwabe, W.L. & Chiesa, J. (1997). *Mandatory Minimum Drug Sentences: Throwing Away the Key or the Taxpayers' Money?* MR-827-DPRC, Santa Monica: RAND.

Centers for Disease Control and Prevention (CDC). Measuring Healthy Days. Atlanta (GA): CDC; November 2000.

Collins DJ and Lapsley HM (2002). *Counting the Cost: Estimates of the Social Costs of Drug Abuse in Australia in 1998-1999*. National Drug Strategy Monograph Series No. 49, Canberra, Australia: Commonwealth Department of Health and Ageing.

Daley M, Shepard DS, and D Bury-Maynard (2005) "Changes in Quality of Life for Pregnant Women in Substance User Treatment: Developing a Quality of Life Index for the Addictions" *Substance Use & Misuse* 40(3): 375-394.

Drummond, M.F., O'Brien, B., Stoddart, G.L., & Torrance, G.W. (1997). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press.

Fenoglio P, Parel V and Kopp, P. (2003). "The Social Cost of Alcohol, Tobacco and Illicit Drugs in France, 1997" *European Addiction Research* 9: 19-28.

French MT, Salome HJ, Sindelar JL and AT McClellan. Benefit-Cost Analysis of Addiction Treatment: Methodological Guidelines and Empirical Application Using the DATCAP and ASI. *Health Services Research* 2002, 37:2 433-455.

Gold, M.R., Siegel, J.E., Russell, L.B., & Weinstein, M.C. 1996. *Cost-Effectiveness in Health and Medicine*. Oxford: Oxford University Press.

Humphreys, Keith, Todd H. Wagner, and Mistry Gage (2008). Is it in medical centers' self-interest to provide substance use disorder treatment?: A cost-consequence analysis in a national health care system. Paper presented at the ISSDP Conference in Lisbon, April 4, 2008.

Jofre-Bonet M. and J.L. Sindelar. Creating an Aggregate Outcome Index: Cost-effectiveness Analysis of Substance Abuse Treatment. *The Journal of Behavioral Health Services & Research* 31(3): 229-241.

Kaplan, Edward H. "A circulation theory of needle exchange", *Operations Research*, vol. 43, No. 4 (1995), pp. 558-569.

Kaplan, Edward H. and Mira Johri. 2000. "Treatment on Demand: An Operational Model," *Health Care Management Science*, Vol. 3, pp.171-183.

Keeler, E., and S. Cretin, Discounting of Life-Saving and Other Nonmonetary Effects, *Management Science*, 29:300-306, March 1983.

Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. *British Medical Journal*. 1998;316:736-41.

Kleiman, Mark AR. (1992) *Against Excess: Drug Policy for Results*. Basic Books, New York.

Kleiman, Mark A.R. and Steven N. Teles (2006) Market and Non-Market Failures. In Michael Moran, Martin Rein, and Robert E. Goodin (Eds.) *The Oxford Handbook of Public Policy*, Oxford: Oxford University Press.

Kuziemko, Ilyana and Steven D. Levitt (2001) An Empirical Analysis of Imprisoning Drug Offenders, NBER Working Paper No. 8489.

MacCoun, Robert and Peter Reuter. 2001. *Drug War Heresies* New York, Cambridge University Press.

Miller, Ted R., Mark A. Cohen, and Brian Wiersema. 1996. *Victim Costs and Consequences: A New Look*. National Institute of Justice Report NCJ 155282, Washington, DC.

Moran, Michael, Martin Rein, and Robert E. Goodin (Eds.) (2006) *The Oxford Handbook of Public Policy*, Oxford: Oxford University Press.

Office of National Drug Control Policy (2004). The Economic Costs of Drug Abuse in the United States, 1992-2002. Washington, DC: Executive Office of the President (Publication No. 207303).

Perez, P. Dray, A., Ritter, A., Dietze, P., Moore, T. & Mazerolle, L. (2006) SimDrug: A Multi-Agent System Tackling the Complexity of Illicit Drug Markets in Australia. In P. Perez and D. Batten (Eds) *Complex Science for a Complex World - Exploring Human Ecosystems with Agents*. Canberra: ANU E press, pp. 193-224.

Pollack, H. A. "Can we protect drug users from hepatitis C?", *Journal of Policy Analysis and Management*, vol. 20, No. 2 (2001), pp. 358-364.

Pollack, Harold. (2002) "Methadone Maintenance as HIV Prevention: Cost-Effectiveness Analysis," in *Quantitative Evaluation of HIV Prevention Programs*, E.H. Kaplan and R. Brookmeyer, eds., Yale University Press, New Haven, CT.

Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001; 33:337-43.

Reuter, Peter and Mark Kleiman "Risks and Prices: An Economic Analysis of Drug Enforcement" in Morris and Tonry (eds.) *Crime and Justice: An Annual Review of Research* Volume 7, 1986, pp.289-340.

Rice Dorothy P. (1967). Estimating the costs of illness. *American Journal of Public Health* 57:424-440.

Single E, Collins D, Easton B, Harwood H, Lapsley H, Kopp P and Wilson E. (2003). *International Guidelines for Estimating the Costs of Substance Abuse*, 2nd Edition, Geneva, Switzerland, World Health Organization.

Stokey, Edith and Richard Zeckhauser (1978) *A Primer for Policy Analysis*. New York: Norton.

Viscusi, Kip W. and Joseph E Aldy (2003) “The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World” National Bureau of Economic Research Working Paper #9487 {National Bureau of Economic Research, Cambridge}.

Weimer, David L. and Aidan R. Vining (2005) *Policy Analysis: Concepts and Practice*. Upper Saddle River: Prentice Hall.

Zarkin Gary A; Dunlap, Laura J; Hicks Katherine A and Daniel Mamo (2005) “Benefits and costs of methadone treatment: results from a lifetime simulation model” *Health Economics* 14: 1133-1150.