The research described in this report was sponsored by the Office of National Drug Control Policy, by the United States Army, and by RAND's Drug Policy Research Center with funding from The Ford Foundation.

ISBN: 0-8330-1553-2

Published 1994 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
To order RAND documents or to obtain additional information, contact Distribution Services: Telephone: (310) 451-7002; Fax: (310) 451-6915; Internet: order@rand.org.

RAND

Modeling the Demand for Cocaine

Susan S. Everingham
C. Peter Rydell

Prepared for the Office of
National Drug Control Policy
United States Army

DRUG POLICY RESEARCH CENTER

PREFACE

This report documents the development of a model of the demand for cocaine that was fit to 20 years of data on the current cocaine epidemic in the United States. It also describes the analysis performed, including the estimation of incidence, prevalence, cohort retention, and consumption. The impetus for the model's development was a parallel RAND analysis of cocaine-control programs (see *Controlling Cocaine: Supply Versus Demand Programs*, C. Peter Rydell and Susan S. Everingham, MR-331-ONDCP/A/DPRC, 1994), of which this analysis is a key component. However, the model of cocaine demand is useful in its own right, leading to new insights on the nature of the cocaine problem.

The work reported here was sponsored by the Office of National Drug Control Policy, the U.S. Army, RAND's Drug Policy Research Center (DPRC) with funding from The Ford Foundation, and RAND's Social Policy Department. The research was jointly carried out within three RAND entities: the DPRC, the National Defense Research Institute (NDRI), and the Strategy and Doctrine Program of the Arroyo Center. NDRI is a federally funded research and development center that supports the Office of the Secretary of Defense, the Joint Staff, and the defense agencies. The Arroyo Center is the U.S. Army's federally funded research and development center.

CONTENTS

FIGURES

TABLES

SUMMARY

Although the status of the "war against cocaine" was still being debated just a few years ago, now it is generally understood that even though the overall number of cocaine users is decreasing, the proportion of those users that are the so-called heavy users is increasing. To elevate the policy debate to a new level, a more precise and quantitative understanding of these trends is required. Toward the goal of designing effective drug control policy, we created a model of how the demand (i.e., the number of users together with how much those users consume) for cocaine changes over time that incorporates available data and interprets them. Specifically, demand is determined by a two-state Markovian model of the user flows that has been fitted to 20 years of historical data on cocaine usage derived from the National Household Survey of Drug Abuse (NHSDA) and other sources.

The Markovian approach to modeling prevalence (the number of people who use drugs) in this analysis can usefully be distinguished from purely statistical techniques such as multiple capture, Poisson estimation, and synthetic estimation, and from elaborate behavioral models such as the system dynamics models. By a Markovian model, we mean one that incorporates one or more states and the transition parameters that determine the flows between those states. We have adopted a two-state, four-parameter model because it supports the most important behavioral distinction, that between light and heavy use, without encumbering the model with unnecessary detail. The four parameters governing transition flows are selected to match the historical data.

Prevalence is a primary indicator of the extent of the illicit drug problem. The principal survey instrument for estimating drug-use prevalence in the United States is, and has been for the last two decades, the NHSDA sponsored by the National Institute on Drug Abuse (NIDA). As its name indicates, the NHSDA reports drug usage among people living in households in the United States. This sampled population includes the vast majority of people 12 and older, but it overlooks some segments of the U.S. population that may include a substantial proportion of drug users, such as the incarcerated and the transient homeless. The prevalence estimates used to establish the model parameters were based upon the NHSDA estimates of the prevalence of cocaine use among the household population supplemented by estimates of cocaine use among the incarcerated and the homeless. The overall prevalence estimates obtained for NHSDA survey years are shown in Figure S.1.

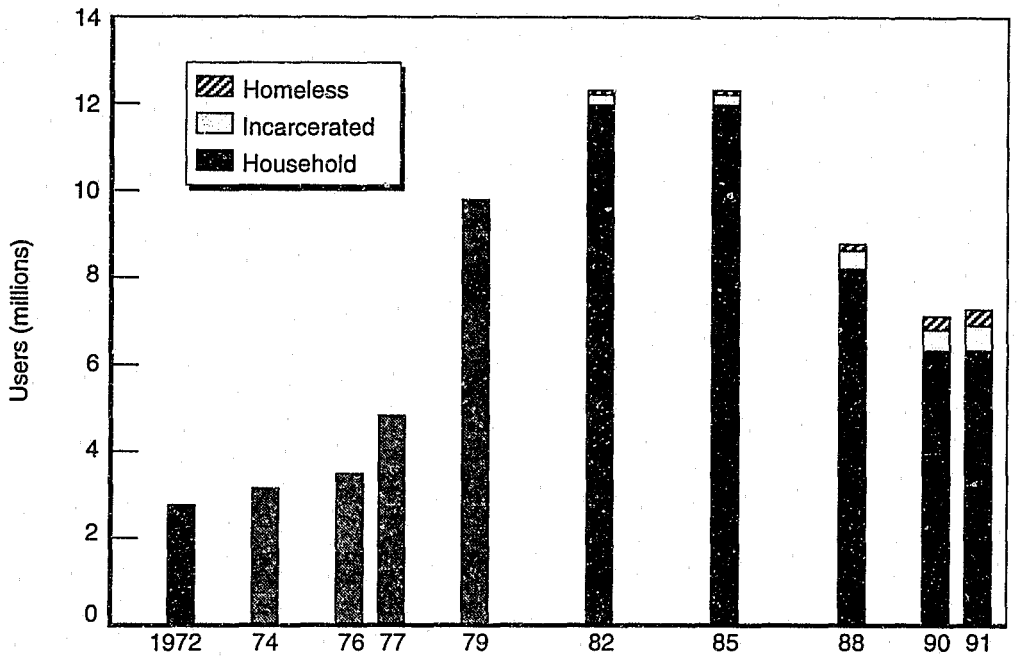

Figure S.1—Overall Prevalence of Cocaine Users in the United States

As a modeling convenience, users were separated into just two categories: light users and heavy users. (Modeling the entire spectrum is neither practical nor necessary, and modeling a single average user is insufficient.) For this model, the distinction between light and heavy use was based simply upon frequency of use. People who said they used at least weekly (or several times a month) were defined as heavy users, and the rest were light users. NHSDA information was used to estimate that the average heavy user consumes eight times as much cocaine as does the average light user.

The Markovian model is required to fit (1) the overall prevalence data; (2) the fraction of all users that were heavy users in 1985, 1988, and 1990; and (3) the fraction of a cohort of initiates that are still using drugs ten years later, i.e., the ten-year cohort retention rate. The incidence (the number of people who initiate drug use) into light cocaine use, which has varied greatly over the years, is an input to the model. (Consequently, the model cannot predict future prevalence; it can only project prevalence given a hypothetical incidence scenario.) The fitting procedure is essentially an exhaustive search of the four-dimensional parameter space. The goodness-of-fit is demonstrated in Figures S.2, S.3, and S.4.

The model demonstrates that the fraction of all cocaine users that are heavy users has varied greatly over time (implying overall prevalence is an incomplete measure of the cocaine epidemic), and that peak heavy usage followed peak incidence (which occurred around 1980) by about ten years (see Figure S.5). Consequently, the effect on heavy cocaine usage of government programs that reduce incidence (such as pre-

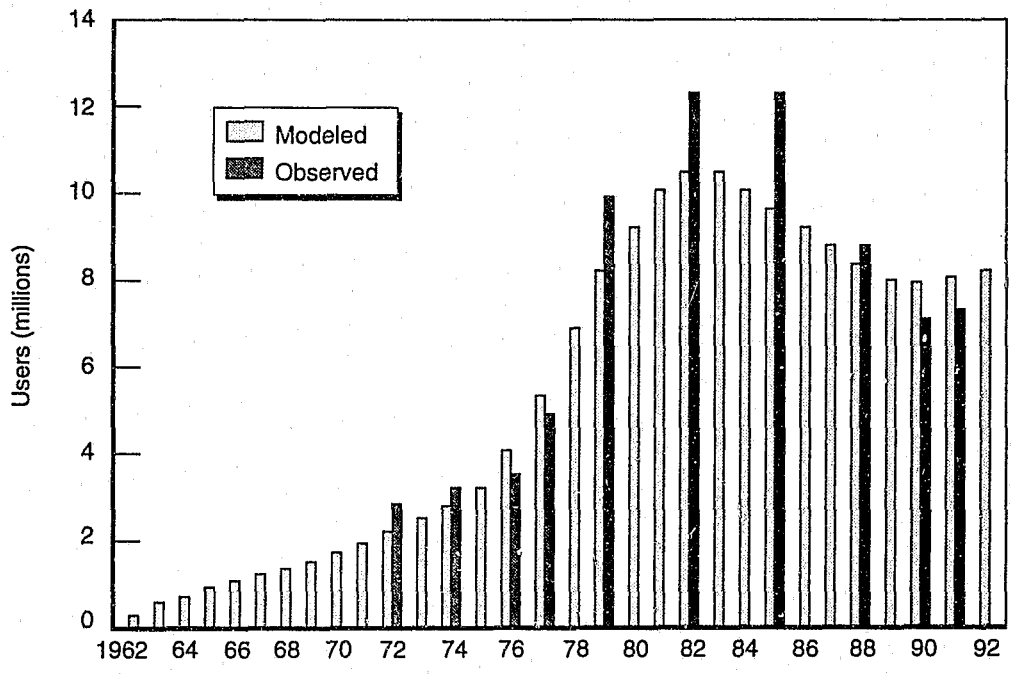

Figure S.2—Overall Prevalence of Cocaine Users: Modeled vs. Observed

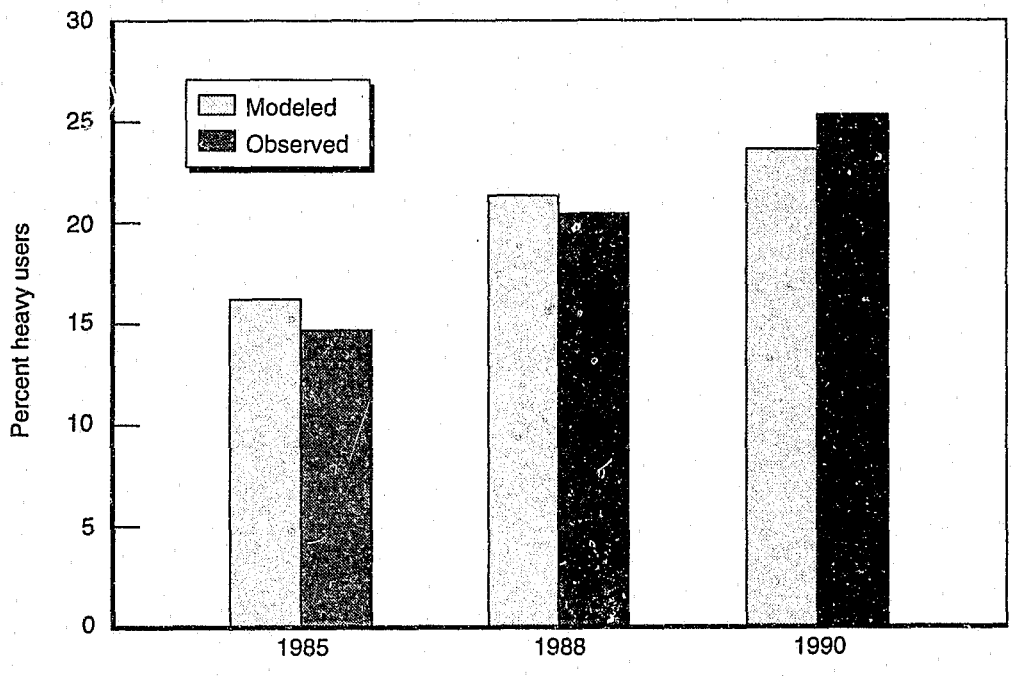

Figure S.3—Percentage of Users That Are Heavy Users: Modeled vs. Observed

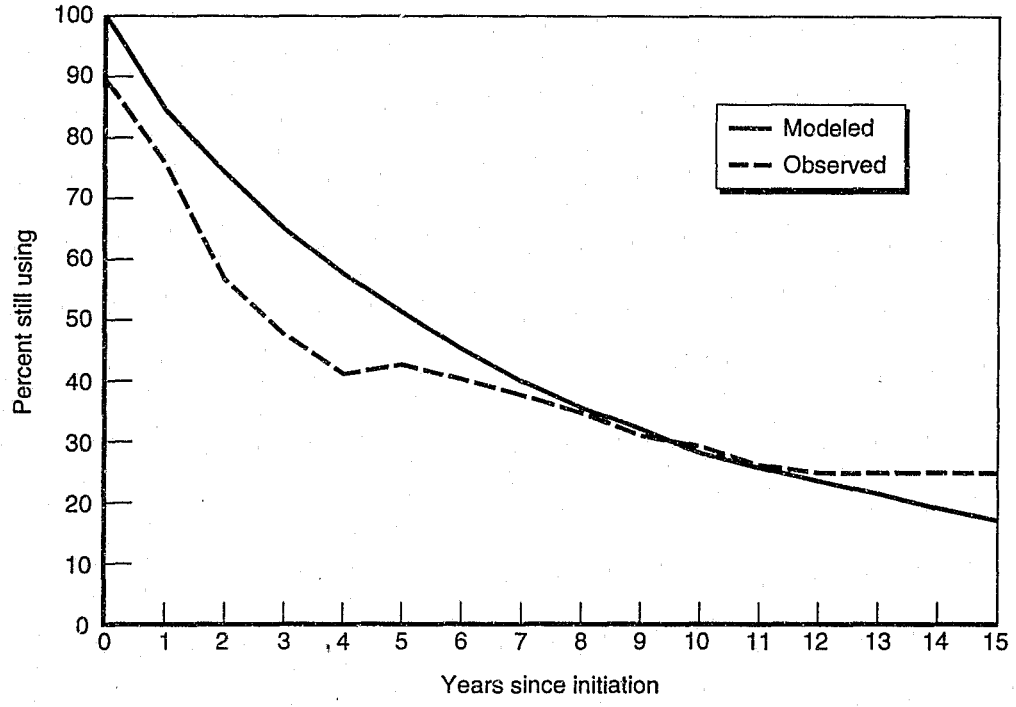

Figure S.4—Cohort Retention: Modeled vs. Observed

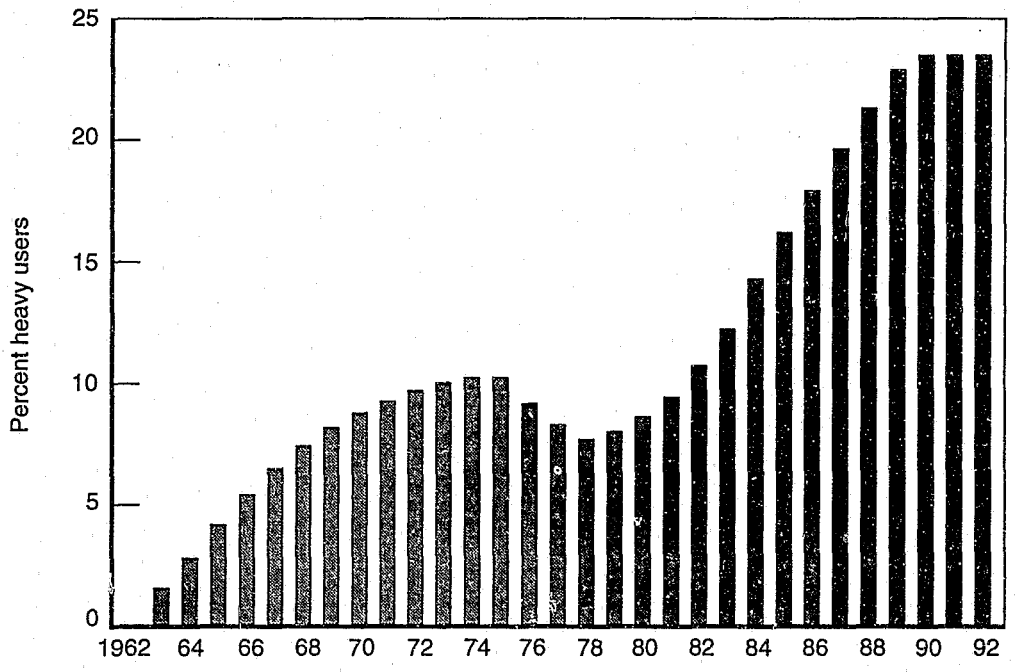

Figure S.5—Modeled Percentage of Users That Are Heavy Users: Variation over Time

vention programs) will only be realized many years later, and part of the effectiveness of local law enforcement programs and other programs that influence drug use in multiple ways (affecting incidence, flow rates, and the consumption rates of current users) also will be delayed. The fact that the various control programs focus upon different aspects of drug use (prevention on incidence, treatment on heavy usage, etc.) means that some strategies may be most appropriate for specific stages of the epidemic.

Figure S.6, a graph of modeled prevalence over time, reveals the underlying contributions to the prevalence estimates by both light and heavy users. The *overall* prevalence curve has characterized the course of the cocaine epidemic in the eyes of many policymakers. But while both overall and light-user prevalence have recently declined and leveled off, the number of heavy users continues to increase.

In contrast to prevalence's overall decline during the past decade, consumption has merely leveled off (see Figure S.7). And even if overall prevalence continues to decline, large amounts of cocaine will still be consumed in the United States because more and more of the remaining users will be heavy users. Given this increasing prevalence of heavy users and its effect on total cocaine consumption, the bottom line is that the "war against cocaine" has by no means been "won."

Although the model cannot predict incidence, it can project the course of the cocaine-use epidemic given any hypothetical incidence scenario. The value of such projections lies in the fact that they bound the analysis in a useful way. Figures S.8 and S.9 plot 15-year projections of prevalence and consumption, respectively, assuming that incidence remains constant at about one million new users per year. The graphs imply that constant incidence, even at the current low level, will result in an increase in both prevalence and consumption.

Assuming (optimistically and probably quite unrealistically) that incidence is reduced to zero and does not resurge, the maximum effect that reduced incidence can have on the future course of the cocaine epidemic can be estimated. From Figure S.10, we see that prevalence is reduced to about two million cocaine users in 15 years. But most of those users are heavy users, so the decrease in consumption is not nearly as dramatic: in 15 years, consumption is only halved (see Figure S.11). Thus, even in the absence of incidence it will take about 30 years for the current epidemic to (nearly) disappear, unless the flow rates out of cocaine use increase.

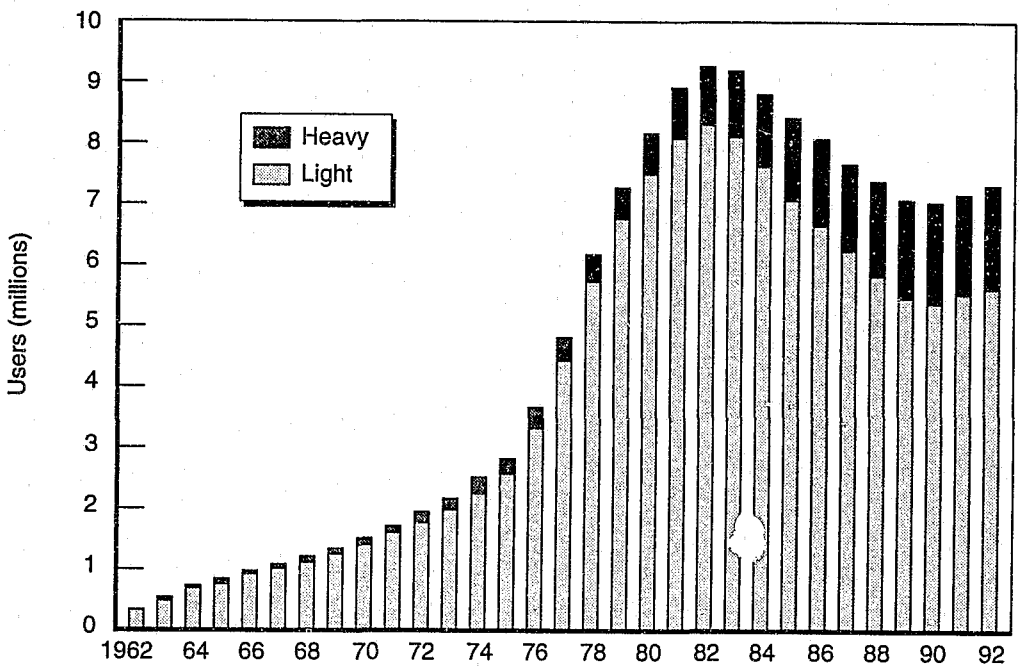

Figure S.6—Modeled Prevalence: Heavy vs. Light Users

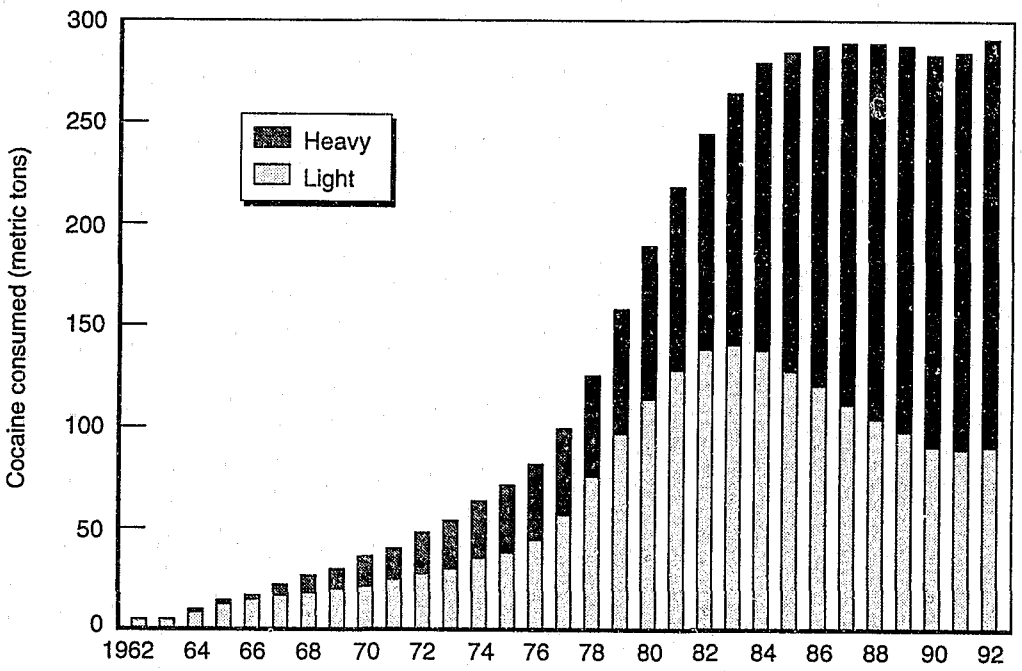

Figure S.7—Modeled Consumption: Heavy vs. Light Users

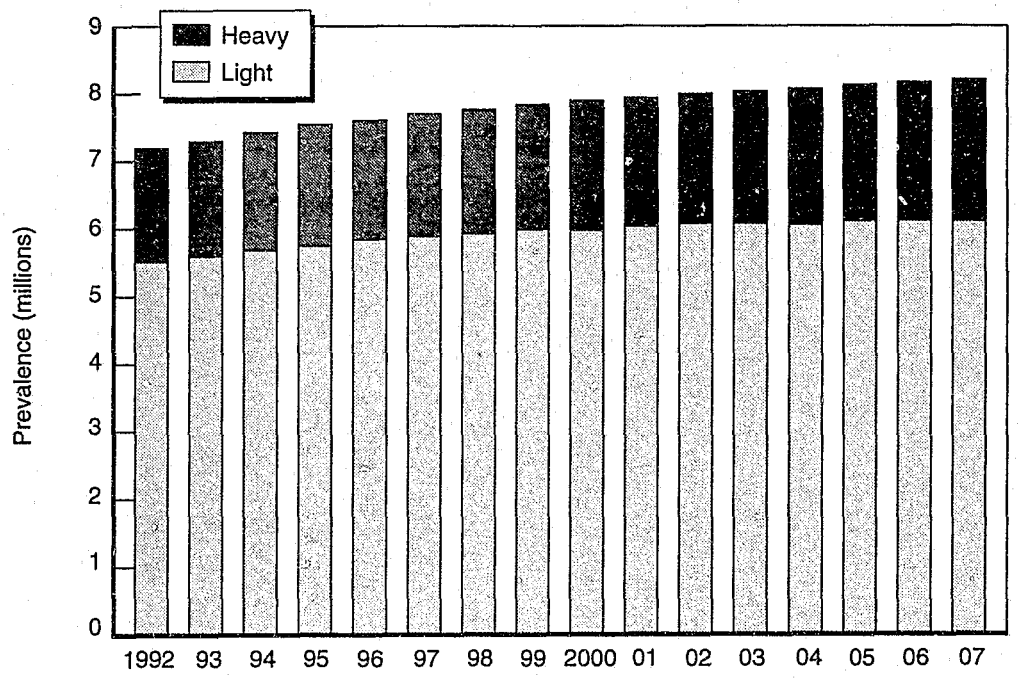

Figure S.8—Prevalence Assuming Constant Incidence

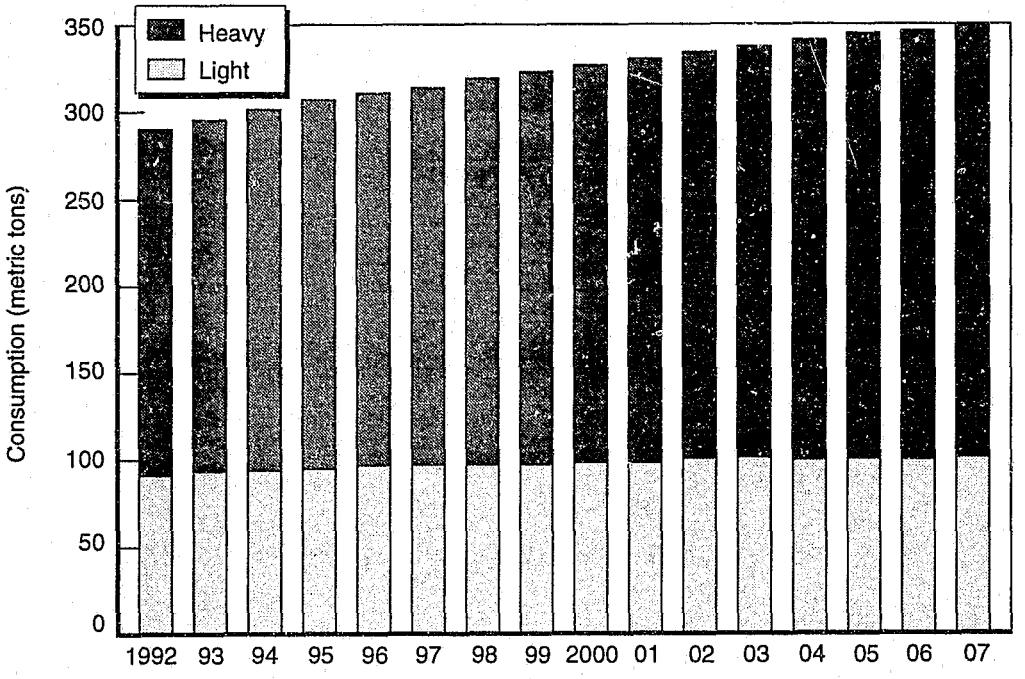

Figure S.9—Consumption Assuming Constant Incidence

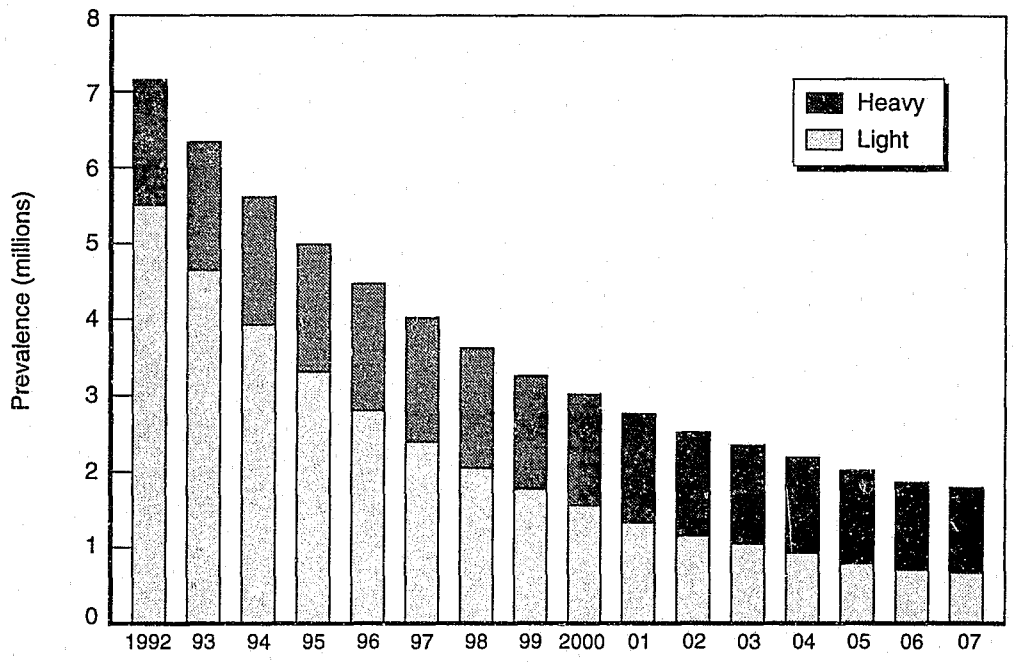

Figure S.10—Prevalence Assuming Zero Incidence

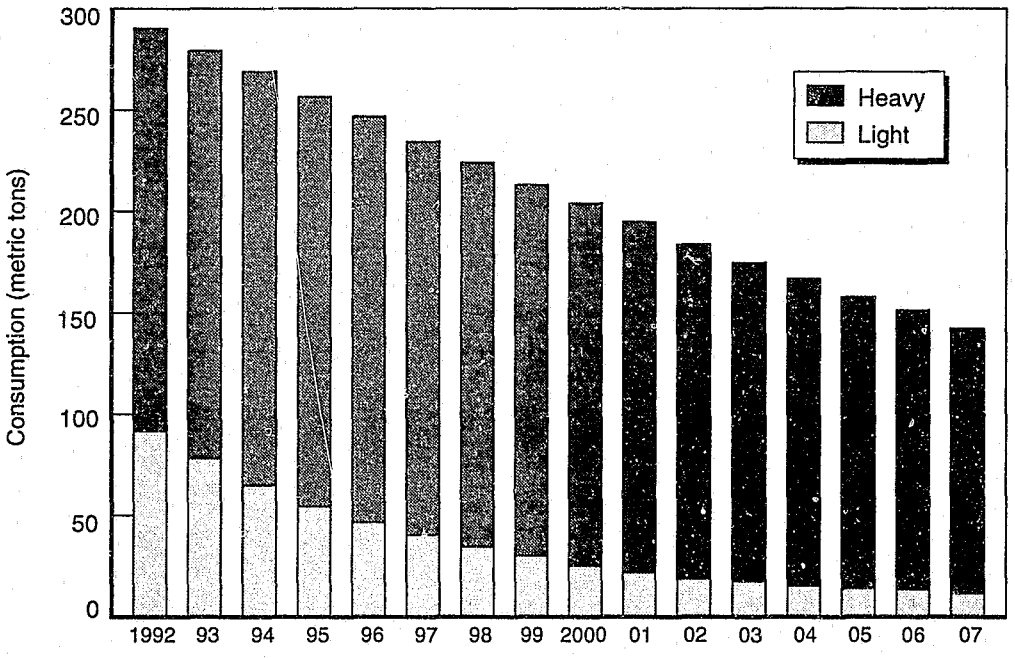

Figure S.11—Consumption Assuming Zero Incidence

ACKNOWLEDGMENTS

The contributions and helpful suggestions of numerous people, both at RAND and elsewhere, proved to be of integral importance to this work and are gratefully acknowledged. In particular, we would like to thank the current co-directors of RAND's Drug Policy Research Center, Audrey Burnam and Jonathan Caulkins, and former co-director Peter Reuter for their heartening support and advice. The insightful reactions to a work-in-progress briefing at the University of California, Los Angeles Drug Abuse Research Center were very useful. Bonnie Dombey-Moore, Patricia Ebener, and, especially, Daniel McCaffrey contributed in a myriad of ways, without which we could not have accomplished our goal.

Chapter One

INTRODUCTION

OBJECTIVE

Have the problems with cocaine use in the United States been getting worse or better? Until recently (and occasionally still today), there seemed to be no satisfactory resolution to the debate. Some pointed to the declining estimates of use among the household population and concluded that the situation was improving. Others pointed to indicator data, such as the number of hospital emergency room mentions of cocaine,[1] and asserted that the situation could only be degenerating. The debate was muddled because people failed to recognize a simple fact: as the cocaine epidemic evolves, different measures of its severity are affected in different ways.

The extent of the cocaine problem in the United States can be measured in a number of different ways: number of users, amount of cocaine consumed, number of people requiring treatment to desist in cocaine use, how often hospitalization is related to cocaine use, societal cost of cocaine use, and so on. Various instruments exist for estimating these quantities. Perhaps the most generally known and used is the National Household Survey of Drug Abuse (NHSDA) sponsored by the National Institute on Drug Abuse (NIDA), which measures the prevalence of cocaine use (i.e., the number of cocaine users) among the U.S. household population. (It is this prevalence data that supports the belief that the cocaine problem is decreasing.[2]) Unfortunately, the data produced by these various instruments are often incomplete, erratic, and contradictory. Thus, integrating the data to produce an overall picture of the cocaine problem in the United States that can sufficiently support decisions about drug control policy is a difficult task, as evidenced by the aforementioned debate.

It is now generally understood that even though the overall number of cocaine users is decreasing, more and more of the users that persist are either addicted users or serious abusers, the so-called "heavy" users. To elevate the debate to a new level, we must now understand more precisely the magnitude of these trends. For effective policy analysis, we must be able to test hypotheses about such trends; for example, has the decrease in the total number of users led to decreased consumption, or has

[1] Emergency room mentions of drug use are tracked by the Drug Abuse Warning Network (DAWN).

[2] Additional evidence of the decline in cocaine usage is found in the Monitoring the Future surveys of high school seniors.

the increase in the number of heavy users more than offset the decrease in total users, causing consumption to increase?

In designing effective drug control policy, a model of how the demand[3] for cocaine changes over time—i.e., a model that incorporates some of the various available data and interprets them—is a useful integrative tool. This report describes a simple version of such a model and presents an analysis aimed at extending the current qualitative understanding of the cocaine situation by providing quantitative estimates of the trends in cocaine demand. We begin by synthesizing data about "what" has happened in the recent cocaine epidemic; we then go on to explore "how" and "why" by modeling the flow of users into and out of cocaine use. Specifically, demand is determined by a two-state Markovian model of the user flows, implemented on a spreadsheet, that has been fitted to 20 years of historical data on cocaine usage derived from the NHSDA and other sources.

This research on the demand for cocaine complements recently completed research on the supply of cocaine.[4] The combined understanding from these two studies clarifies the ways in which the cocaine epidemic responds to alternative cocaine-control programs—supply-control programs such as interdiction, and demand-control programs such as drug treatment. The research described here also feeds into a broader analysis by providing a baseline model of cocaine-demand dynamics that can be used to measure and compare the effects of policy changes.[5]

In addition to supporting the broader analysis, the model elucidates information that is difficult or impossible to extract or intuit directly from the data sources, and facilitates comparison of those data. It allows exploration of the dynamics of the cocaine epidemic, both the trends and the flows. Moreover, given a hypothetical scenario of incidence (i.e., a specified number of new cocaine users in a given period of time), it can project a course for the cocaine epidemic (the validity of the projection will, of course, depend on the accuracy of the incidence scenario).

BACKGROUND

The Markovian approach to modeling prevalence in this analysis can usefully be distinguished from purely statistical techniques such as multiple capture, Poisson estimation, and synthetic estimation,[6] and from elaborate behavioral models such as the system dynamics models.[7] Compared to the purely statistical methods, which offer only a point estimate of prevalence, our Markovian model has more behavioral content—i.e., flows into and out of use, and consumption rates. Compared to system

[3]Demand combines the number of cocaine users with the amount of cocaine they are consuming.

[4]See Dombey-Moore, Resetar, and Childress (forthcoming).

[5]The broader analysis, in which various drug control policy options are compared, is described in Rydell and Everingham (1994).

[6]See, for example, Rhodes (1993).

[7]See the recent review articles by Hser (1993) and Wickens (1993).

dynamics models, however, our model has less behavioral content.[8] In particular, it does not include the feedback effect of prevalence on incidence. As Musto (1973) points out in his historical analysis of a century of drug use, drug epidemics eventually end when, with time, a new generation becomes sufficiently aware of the dangers to impede the inflow of new users. Unfortunately, after two (or so) generations have passed, awareness of the dangers fades, and incidence can resurge.

Rather than modeling incidence and this feedback effect, our model scripts incidence. That is, incidence estimates determined from historical data are used when fitting the model, and incidence scenarios are used for projecting the future. While not suited for modeling epidemics on the macro scale, our approach is useful for short- and intermediate-range prevalence estimation. It is particularly useful for analysis of an ongoing epidemic, as is currently the case with cocaine.

DATA USED TO FIT MODEL

The NHSDA was the primary instrument used to determine the model parameters. It is an occasional (more recently, annual), extensive survey of drug usage in the United States. The NHSDA focuses on the U.S. household population and therefore misses institutionalized populations (such as the incarcerated) and (until recently) the homeless. The NHSDA estimates were modified by adding estimates of the number of cocaine users among the incarcerated and homeless, and then the model was fit to this composite population estimate. This adjustment is important because drug users are more prevalent among the incarcerated and homeless populations than among the U.S. population in general. It turns out that the sizes of these additional populations are small compared to the household population, so the effect on the overall prevalence estimates presented here is minor. However, since heavy users are overrepresented among these nonhousehold populations, the effect on heavy-user prevalence is more dramatic than that on total prevalence.

LIMITATIONS OF THE MODEL

One shortcoming of this analysis is that the prevalence estimates implied by the NHSDA may be too low, even after adjustment to account for the nonsurveyed populations (the homeless and the incarcerated).[9] Studies have established point estimates of the prevalence of drug usage via other and/or broader means (Rhodes, 1993), but they too are subject to significant uncertainty. For the purposes of this analysis, a series of prevalence estimates over time was required. No existing data other than the NHSDAs are sufficiently consistent over time to serve as the basis for this modeling exercise. Our prevalence estimates could conceivably have been im-

[8]For example, see Levin (1975) and Homer (1990, 1993).

[9]This shortcoming results from possible nonresponse and underreporting biases. Since the NHSDA relies on self-report of drug use, the potential for underreporting bias certainly exists (see, for example, Falck et al., 1992), even though Mieczkowski (1990) found that self-reported drug use is often accurate. Regarding the NHSDAs, the response rate in all but two surveys was good (at least 80 percent), and the 1988 and 1990 surveys (at least) were adjusted to account for nonresponse bias. The magnitude of the bias linked to the pattern of nonresponse in the NHSDAs is likely to be small (Harrison, 1991).

proved, but not without considerable effort. And in any case, uncertainties inherent in the data would have remained to overshadow the improvement.

Obviously, the validity of the model hinges upon the validity of the data on which it is based. Better estimates of prevalence among the homeless and the incarcerated than ours (which are only rough) and better estimates of the history of incidence would no doubt improve the model's validity. However, our findings would be unlikely to change, as sensitivity excursions have demonstrated.

This analysis attempts neither to measure changes in the flow rates nor to explain the forces behind those changes. In particular, when estimating the flow rate parameters in the model of cocaine demand, we did not control for changes over time in the price of cocaine or the availability of treatment. The general trends during the period were that the price of cocaine fell and the availability of treatment increased. These trends tend to have opposite effects on the flow rates (i.e., decreasing price should decrease outflow, whereas increased treatment should increase outflow). In essence, we ignored these dynamic effects and fit the parameters to the "average conditions" in price and treatment over the period modeled (from the early 1960s to the beginning of the 1990s).

We use *cocaine* to mean either crack or powder. This analysis thus covers both but does not distinguish between them. The introduction of crack in the late 1980s may have altered the patterns of cocaine use—for example, crack users may move more rapidly than powder users from casual use to addiction. Understanding such differential effects was beyond the scope of our modeling effort.

REPORT ORGANIZATION AND OVERVIEW

The rest of this report is divided into seven chapters. The first of these, Chapter Two, describes the generic Markovian modeling concept and explains the rationale for the two-state, four-parameter model structure.

Chapter Three then presents in detail the prevalence data to which the model was fit. This model separates users into two categories: light users and heavy users. This approach represents a compromise between modeling the entire spectrum (which is infeasible) and modeling a single average user (which is insufficient). It is consistent with the intuitive belief that heavy users should be viewed and counted differently than light users because of the different social costs associated with heavy cocaine consumption. Heavy users are defined to be people who use cocaine at least weekly. In addition to the overall prevalence numbers, the fraction of all users that are heavy users and the relative consumption rates of light and heavy users are discussed in Chapter Three.

A cohort retention rate gives the fraction of a cohort of initiates still using the drug after a given period of time. We calculate cohort retention rates from NHSDA data in Chapter Four; various estimates of annual incidence, which is an input to the model, are described in Chapter Five.

The four unknown parameters of the Markovian model, the flow rates, are determined by the fitting procedure explained in Chapter Six. This procedure requires the model to match (1) the overall prevalence data for the entire course of the current epidemic, (2) the fraction of all users that are heavy users over recent time, and (3) the ten-year cohort retention rate. This analysis determines the *fixed* flow rates that best match the historical data.[10]

Interesting observations about the history of demand that are not directly evident from the data alone but are highlighted by the model are discussed in Chapter Seven. Prevalence projections based on hypothetical incidence scenarios are presented in Chapter Eight.

---

[10]The dynamic nature of the modeled system deserves emphasis: along with prevalence and incidence, flow rates and consumption rates also vary with time.

Chapter Two

MARKOVIAN MODELS OF DEMAND

By a Markovian model, we mean one that incorporates one or more states and the probabilities of transition between them. The transition probabilities depend only upon the existing state of the system. The model can be represented by a simple system of (possibly nonhomogeneous) linear difference equations: $Q(t_i) - Q(t_{i-1}) = A*Q(t_{i-1}) + F(t_i)$, where $\{t_i\}$ is the discretized time variable, Q is some vector quantity representing the states of the system, A is the matrix of transition probabilities, and F is the optional nonhomogeneity known as the forcing function. The corresponding differential equation is $dQ(t)/dt = A*Q(t) + F(t)$.

The simplest Markovian model is one that includes only one state, in which Q, A, and F are scalar quantities. From elementary calculus, one recognizes that the solution to the corresponding homogeneous differential equation (i.e., $F(t)$ is identically zero) is $Q(t) = Q(0)*\exp(A*t)$. This system is either constant, exponentially growing, or exponentially decaying. The solution to the homogeneous, higher-dimensional system is also straightforward. It can be represented by the same solution equation if $\exp(A*t)$ for a matrix A is defined appropriately. The geometry of the solution is again one of only a handful of possibilities. The presence of a forcing function (the nonhomogeneity $F(t)$) greatly complicates the geometry, regardless of the dimension of the system.

The goal of this research was to develop a dynamic (i.e., time-dependent) model of the number of cocaine users in order to better understand the flow of users into and out of drug use. For this application, we considered the population of non-users to be unlimited in size, so non-use is not a quantified state in the model. Moreover, we assumed that a flow is only dependent on the magnitude of the source; that is, the flow from state 1 to state 2 is proportional to the size of the state 1 pool only.[1] Therefore, the flow of people from non-use to use, the incidence, is quantified by a time-dependent forcing function. The time step of the model, consistent with the available data, is one year.

[1]More complicated, usually nonlinear, models, in which the flow from a source is a function of the sizes of pools other than the source, are commonly hypothesized in many applications, including epidemiology. For example, one could hypothesize that the flow into drug use is proportional to the number of current users, since current users are the agents of "infection." The development of such models would be of significant theoretical and practical interest, but it is unclear if enough data exist to support their validation. As such, our approach was to create a simple but credible model that can later be further developed, which is the only prudent way to develop a model of a very complicated system.

There are several such models, ranging in complexity, that are applicable. The simplest is the one-state model diagrammed in Figure 2.1. The year is represented by *t*, the annual incidence is represented by *I(t)*, the number of users is represented by *U(t)*, and the flow of users to non-use is represented by *a*U(t)*. As discussed in Chapter One, the values of the unknown parameters (the transition probabilities) that make the model best fit the available data are determined by the analysis; this model has only one unknown parameter, the transition probability *a*. This model, however, was deemed too simplistic for two reasons. First, it does not distinguish between light and heavy users, and analysis of NHSDA data supports the importance of this distinction (see Chapter Three). Second, it could not be well fit to the prevalence and cohort retention data (i.e., no value of *a* provided for an adequate fit of the data).

If users are divided into two groups—light and heavy users—that are counted separately, then a two-state, four-parameter model is generated (see Figure 2.2). This model could be further refined in one of two ways, both of which increase the number of unknown parameters that must be fit to the data and hence add to the complexity of the model and the fitting procedure.

The first possible refinement is to divide the users into more than two groups (such as light, medium, and heavy users). This option was deemed superfluous and not supportable by the available data. The second possible refinement is to have users flow into "previous user" pools instead of returning to the non-user pool. This option, not covered in our analysis, does have merits (such as permitting a distinction to be made between incidence and relapse) that suggest it should be further explored. The reality is that the dynamics of cocaine use could be represented by many different such Markovian models. Of course, fitting to these more complex models, which have more than four unknown parameters, would be significantly more difficult.[2]

We adopted the two-state, four-parameter model for this analysis because it is complex to a necessary and sufficient degree. It supports the most important behavioral

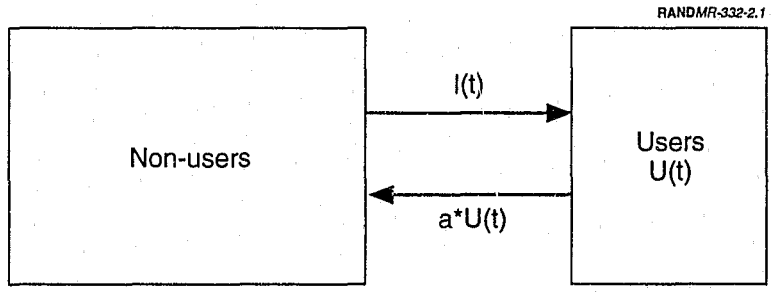

Figure 2.1—A One-State Markovian Model

[2]Dividing users into only two groups is, indisputably, a modeling convenience, since users exhibit not just two, but rather a wide variety of behavior patterns. However, model building always requires a compromise between simplicity and detail, the main driver of which is the character of the supporting data.

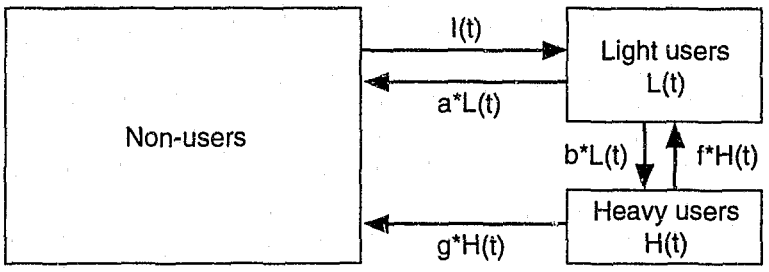

Figure 2.2—A Two-State Markovian Model

distinction, that between light and heavy use, without encumbering the model with unneeded detail and without requiring excessive extrapolation of the available data. The data to which the model was fit are described in Chapters Three and Four.

In our analysis, people are considered either non-users of cocaine (a group assumed to be unlimited in size), light users, or heavy users.[3] New users enter only the light-user pool. The flow of non-users to light use, which is the incidence *I*, is a scripted input to the model. That is, the counts of light users are adjusted each year (the time step of the model) by an (external) estimate of the number of new users.[4]

Some light users flow on to heavy use, but most flow out of the light-user pool into non-use, reflecting the natural tendency of most initiates to quit using cocaine. Heavy users flow back to light use or out of cocaine use. *L(t)* and *H(t)* represent the time-dependent (i.e., year-dependent) numbers of light and heavy users, respectively. The fraction of light users that flow out of cocaine use each year is denoted by *a*, the fraction of light users that flow on to heavy use is denoted by *b*, the fraction of heavy users that flow back to light use is denoted by *f*, and the fraction of heavy users that flow out of cocaine use is denoted by *g*. These four flow rates (also known as transition probabilities) are the fractions of people who flow from the various states during a given year. They are the unknown parameters that must be chosen to fit the historical data.[5]

The model can be represented by a system of two linear, nonhomogeneous difference equations:

$$L(t) - L(t-1) = -(a+b)^*L(t-1) + f^*H(t-1) + I(t)$$

$$H(t) - H(t-1) = -(f+g)^*H(t-1) + b^*L(t-1)$$

[3]This Markovian model is called a two-state model because two pools (light users and heavy users) are tracked in size. Changes in the size of the non-user pool are not tracked.

[4]Chapter Five presents the annual incidence estimates assumed for the model. They count only new users, i.e., people who have used cocaine in the past year for the first time. Cocaine users who quit for a number of years and then relapse are not explicitly modeled.

[5]As discussed in Chapter One, even though flow rates probably vary with time, our analysis determines the fixed flow rates that best match the historical data.

Chapter Three

ESTIMATES OF THE PREVALENCE OF COCAINE USE OVER TIME

A primary indicator of the extent of the illicit drug problem is prevalence, or the number of people who use drugs. The importance of this indicator is highlighted by the degree to which the government's policymakers rely on various prevalence estimates, especially those derived from the NHSDA, to measure the drug problem. In fact, six out of eleven of the goals detailed in the *National Drug Control Strategy* (Office of National Drug Control Policy, 1992) are based on prevalence. Although prevalence is not the only relevant indicator,[1] it is clearly an important element of the overall picture. Accordingly, the prevalence of cocaine use in the United States was one of the pieces of information used to determine the parameters of our model.

The principal survey instrument for estimating drug-use prevalence in the United States is, and has been for the last two decades, the NHSDA, which is sponsored by NIDA. The NHSDA reports drug usage among people aged 12 and older who are living in households in the United States.[2] Although the sampled population includes the vast majority[3] of people twelve and older living in the United States, it omits some segments of the U.S. population that may include a substantial proportion of drug users, such as the incarcerated and the transient homeless. The prevalence estimates we used to establish the model parameters were based upon the NHSDA-derived prevalence estimates supplemented by estimates of cocaine use among the incarcerated and homeless.

ESTIMATES OF PREVALENCE DERIVED FROM THE NATIONAL HOUSEHOLD SURVEY OF DRUG ABUSE

The NHSD, which has been administered intermittently since 1971[4] and annually since 1990, selects a random sample of the entire population of the United States

[1]Others include (1) the estimated need for drug addiction treatment, as is championed by and estimated in a report by the Institute of Medicine (Gerstein and Harwood, 1990); (2) the number of drug-related emergency room episodes, which is compiled by DAWN; and (3) the attitudes of high school students toward drugs, which are monitored in an annual survey administered to the nation's high school seniors that is known as both Monitoring the Future (MTF) and the High School Senior Survey (HSSS). (See Ebener, Feldman, and Fitzgerald, 1993, for a list of drug-related databases.)

[2]The 1991 survey included, for the first time, some nonhousehold populations (described below).

[3]More than 99 percent, according to the U.S. Department of Health and Human Services (1991).

[4]Respondents were first asked about cocaine use in the 1972 NHSDA.

living in households (and, since 1991, living in some group quarters, such as civilians in military installations, students in college dormitories, and homeless in shelters). For each of several illicit and licit drugs, each respondent is asked (utilizing procedures designed to assure confidentiality) about any lifetime use, use during the past year, and use during the past month. Some respondents are also asked about their drug-use behaviors (such as frequency, quantity, age at first use).

Table 3.1 shows the estimated sizes of the populations at risk (i.e., the surveyed populations). It also shows the percentages of those populations that reported lifetime, past-year, or past-month cocaine use for each of the ten surveys conducted from 1972 to 1991 for which results regarding cocaine use were available.

These data, translated into the number of people reporting lifetime, past-year, and past-month cocaine use, are plotted in Figure 3.1.[5] That the past-year and past-month curves were (until 1991) decreasing has been considered evidence that the nation's cocaine problem was becoming less severe.[6] The recent leveling-off in the decline in past-year and past-month use has been recognized as a deceleration in progress against drug use. This deceleration has been credited to the fact that chronic, addictive drug use is much harder to combat than is casual, experimental use; progress is expected to become increasingly more difficult as a greater percentage of the users become chronic, addicted drug users (Office of National Drug Control Policy, 1992). The prevalence of drug usage thus may be an insufficient measure of the extent of the drug problem, a possibility that is further explored in Chapter Seven.

Table 3.1

Estimates of Populations at Risk and Percentage of Population Reporting Cocaine Use

	Population at Risk (millions)			Percentage of Population Reporting Lifetime/Past-Year/Past-Month Cocaine Use[a]		
Survey Year	Ages 12–17	Ages 18–25	Ages 26 and up	Ages 12–17	Ages 18–25	Ages 26 and up
1972	24.662	27.978	107.584	1.5/1.5/0.6	9.1/0.0/0.0	1.6/0.0/0.0
1974	25.047	30.158	112.422	3.6/2.7/1.0	12.7/8.1/3.1	0.9/0.0/0.0
1976	24.797	31.516	116.223	3.4/2.3/1.0	13.4/7.0/2.0	1.6/0.6/0.0
1977	24.938	30.553	117.266	4.0/2.6/0.8	19.1/10.2/3.7	2.6/0.9/0.0
1979	23.419	31.985	123.954	5.4/4.2/1.4	27.5/19.6/9.3	4.3/2.0/0.9
1982	23.304	33.072	126.105	6.5/4.1/1.6	28.3/18.8/6.8	8.5/3.8/1.2
1985	21.640	32.490	136.660	4.9/4.0/1.5	25.2/16.3/7.6	9.5/4.2/2.0
1988	20.250	29.687	148.409	3.4/2.9/1.1	19.7/12.1/4.5	9.9/2.7/0.9
1990	19.978	29.020	152.189	2.6/2.2/0.6	19.4/7.5/2.2	10.9/2.4/0.6
1991	20.144	28.496	154.218	2.0/1.5/0.4	18.0/7.7/2.0	12.0/2.5/0.8

SOURCE: NHSDA, various years.

[a]Where NHSDA estimates were unavailable or too low to be of sufficient precision, estimates of 0.0 were used.

[5]The 1988 survey reported fewer lifetime users than did the 1985 survey, which is not possible unless a disproportionate and highly unlikely number of lifetime users died in the interim. This discrepancy was corrected by interpolating adjacent data points.

[6]See, for example, the report by the Office of National Drug Control Policy (1992).

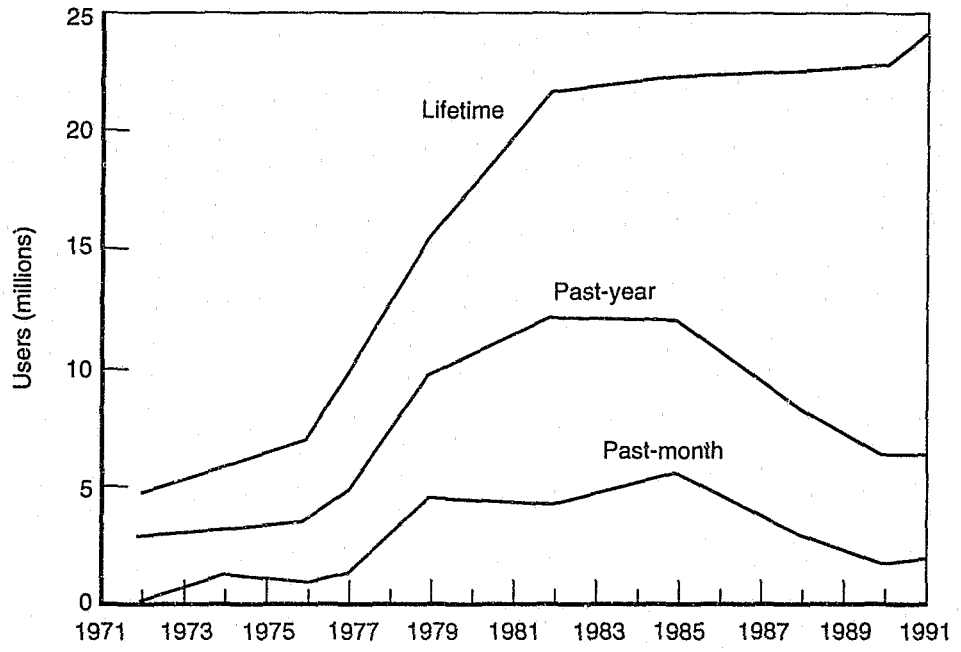

Figure 3.1—Number of People Reporting Lifetime, Past-Year, and Past-Month Cocaine Use

The numbers of people reporting past-year use were the basis for the prevalence estimates used to determine the model parameters. The estimates of the number of lifetime users were utilized to determine incidence in the procedure described in Chapter Five. The number of people reporting past-month use is sometimes regarded as a surrogate for the number of people currently and regularly using cocaine, but was not so regarded in this study. One problem with past-month counts is that they include people who use cocaine infrequently but by chance used it in the past month. Another problem is that regardless of the survey interviewer's assurances of confidentiality, the fact that cocaine consumption is an illegal activity may make some people unwilling to admit to past-month use, even if they will accurately report past-year use.

DEFINITION OF LIGHT AND HEAVY USERS

Like any human behavior, cocaine usage varies across a spectrum. Some people use very little cocaine and only infrequently, some use a large amount daily, and some exhibit just about every behavior in between. As a modeling convenience, users were separated into just two categories: light users and heavy users. Modeling the entire spectrum is neither practical nor necessary, and modeling a single average user is insufficient (as discussed in Chapter Two and further explored in Chapter Seven). The average quantity consumed per user per year has changed substantially over the years because, as shown below, the distribution of user types has changed.

The conditions under which a cocaine user is considered to be a heavy user are not unambiguously defined. A number of criteria—including frequency of use, quantity of cocaine consumed, history of drug use, and the extent of adverse consequences to drug consumption—are all clearly relevant. For example, heavy and light users could be defined by the amount of cocaine consumed by each user. The problem with this approach, however, is that individuals are unlikely to precisely estimate how much they have consumed over a long period of time. They may be able to recollect how much they used the last time, but are unlikely to know how much they used several months ago.[7] Presumably people estimate frequency of use more accurately. So, for this modeling exercise, the definition of light and heavy use was based simply upon frequency of use. The NHSDA asks people who used cocaine in the last year how frequently they used it. People who said they used it at least weekly were defined as heavy users.[8] All other people who had used cocaine in the last year were defined as light users.

Clinicians and researchers commonly divide drug consumption into three levels: use (experimental, occasional, social consumption), abuse (regular, sporadically heavy, intensified consumption), and dependence (compulsive or addictive consumption).[9] While these distinctions are undoubtedly clinically significant, this categorization of users is not easily derived from current prevalence estimating tools. In Gerstein and Harwood (1990), questions in the 1988 NHSDA similar to the World Health Organization's ICD-10 and the American Psychiatric Association's DSM-III-R diagnostics are used to determine the extent of the need for drug treatment. The latter two categories of abuse and dependence together approximately make up the group in need of treatment; this group roughly corresponds to the category of heavy use in our analysis.

Table 3.2 shows the percentage of users in each of the eight frequency categories from the 1990 NHSDA. By our frequency definition, 78 percent of all cocaine users (in 1990) would be considered light users. Notice that the category with the largest percentage of users is that corresponding to least frequent usage. Presumably there is a tendency among users to underreport both frequency and quantity of drug consumption. For this analysis, we assumed there was no significant bias in that underreporting—i.e., that light users underreport to the same degree as heavy users do.

The NHSDA asks cocaine users who responded positively to the past-month use question how much they consumed in that month. Crossing these data with the fre-

---

[7]Some users might be able to estimate how much money they have spent on drugs. But if they share their purchases or if the price of the drug is volatile, total amount of money spent would not translate well into an estimate of usage.

[8]For this analysis, "several times a month" and "at least weekly" were considered equivalent.

[9]The criteria for abuse and dependence are codified in the tenth edition of the *International Statistical Classification of Diseases, Injuries, and Causes of Death* (ICD-10), recently produced by the World Health Organization, and the third revised edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III-R), produced in 1987 by the American Psychiatric Association. Each system offers an array of nine criteria, such as "progressive neglect of alternative pleasures or interests in favor of substance abuse" and "marked tolerance," any three of which trigger a diagnosis of dependence. Abuse is characterized by persistent substance use despite adverse consequences (DSM-III-R) or evidence that the substance causes the user actual psychological or physical harm (ICD-10). (Gerstein and Harwood, 1990.)

Table 3.2

Definition of Light and Heavy Cocaine Use

Category	Reported Frequency of Use During Year	Type of User	Percentage of All Users	Cumulative Percentage
1	1–2 days/year	Light	39.4	39.4
2	3–5 days/year	Light	18.9	58.3
3	Every other month or so	Light	9.6	67.9
4	1–2 times a month	Light	10.2	78.0
5	Several times a month	Heavy	10.2	88.2
6	1–2 days/week	Heavy	6.2	94.4
7	3–6 days/week	Heavy	4.4	98.8
8	Daily (6 days/week or more)	Heavy	1.2	100.0

SOURCE: 1990 NHSDA.

quency data provides an estimate of how much cocaine is consumed by people in each of the eight NHSDA frequency categories. These data from the 1990 NHSDA are displayed in Table 3.3.

The resulting past-30-day consumption by all members of each frequency group is shown in Table 3.4.[10] Seventy-eight percent of all users in 1990 were considered light users by our frequency definition, but that group consumed only about 30 percent of the cocaine. Heavy users, a group that was smaller in number by a factor of

Table 3.3

Reported Number of Past-Month Cocaine Users Distinguished by Frequency and Amount (in thousands)

		Grams Consumed During Past 30 Days[a]						
Category	Past-Month Users	0.125	0.25	0.5	1.0	2.0	3.0	4.0
1	264.0	142.9	29.0	19.6	0.0	0.0	0.0	0.0
2	157.9	58.9	30.3	15.9	0.0	19.0	0.0	0.0
3	124.2	72.6	0.0	1.9	0.0	0.0	0.0	0.0
4	289.6	57.5	44.3	49.5	10.0	103.3	0.0	0.0
5	326.5	60.6	34.6	93.5	12.8	9.8	42.0	40.2
6	198.9	20.6	66.5	0.0	40.2	52.4	0.0	6.1
7	130.3	23.8	0.0	0.0	0.0	22.2	36.0	30.7
8	36.9	0.0	0.0	0.0	2.8	0.0	1.4	31.7
N/A[b]	73.1	32.2	0.0	0.0	4.5	0.0	0.0	9.6

SOURCE: 1990 NHSDA.

[a]These totals include cocaine consumed both as powder and as crack. The conversion factor of 0.1 grams of cocaine per vial of crack was assumed.

[b]No answer provided.

[10]The total amount consumed during the past 30 days by people in each frequency group was determined as follows. The number of past-month users in each frequency group was adjusted up to account for the past-month users who did not respond to the frequency question. Then, the total amount consumed during the past 30 days by all people in a given frequency group was calculated by multiplying the number of people in an amount category by the corresponding amount, summing over all seven amount categories, and adjusting the number upward to account for the past-month users who did not respond to the amount question.

Table 3.4

Estimated Amount of Cocaine Consumed During Past 30 Days by All Frequency Groups

Category	Cumulative Percentage of Users	Percentage of Consumption	Cumulative Percentage of Consumption
1	39.4	3.2	3.2
2	58.3	5.1	8.3
3	67.9	1.1	9.4
4	78.0	18.8	28.2
5	88.2	28.1	56.3
6	94.4	13.3	69.6
7	98.8	21.3	90.9
8	100.0	9.1	100.0

SOURCE: 1990 NHSDA.

almost four, consumed the rest of the cocaine.[11] Simply put, a large group of people used a small fraction of all the cocaine consumed in the United States, and a relatively small group of people used the rest—i.e., the vast majority. This finding is reflected in the concavity of the Lorenz curve, the (smoothed) cumulative percentage of consumption versus the cumulative percentage of users, plotted in Figure 3.2.

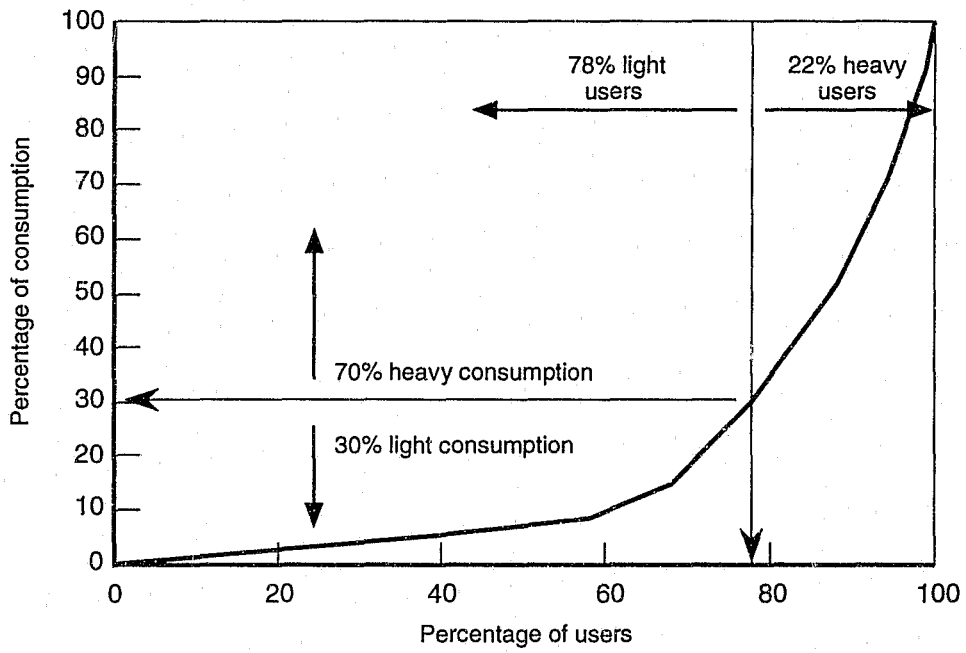

Figure 3.2—Cumulative Percentage of Consumption vs. Cumulative Percentage of Users

[11] These fractions are by no means constant; see below.

This (smoothed) information can also be used to determine that heavy users on the average consume annually eight times as much cocaine as do light users, since (30.0/78.0)/(70.0/22.0) = 1/8 (see Figure 3.3).[12] Note that if the model considered only an average cocaine user, the fact that some users consume significantly more cocaine than others (and thus are perhaps more amenable to treatment) would be lost.

If the NHSDA accurately estimates both the number of users and how much those users consume, it should be possible to estimate the total amount of cocaine consumed by all users (in 1990) by simply multiplying the total consumption, 1585.7 kilograms, by the number of 30-day periods in a year (365/30). This calculation, however, leads to a total of only 19.3 metric tons, which is far less (more than an order of magnitude lower) than the amount estimated by other means to be consumed in the United States.[13] Thus, either the number of past-month users or the amount those users consumed in the past 30 days, or both, must be significantly too low.[14]

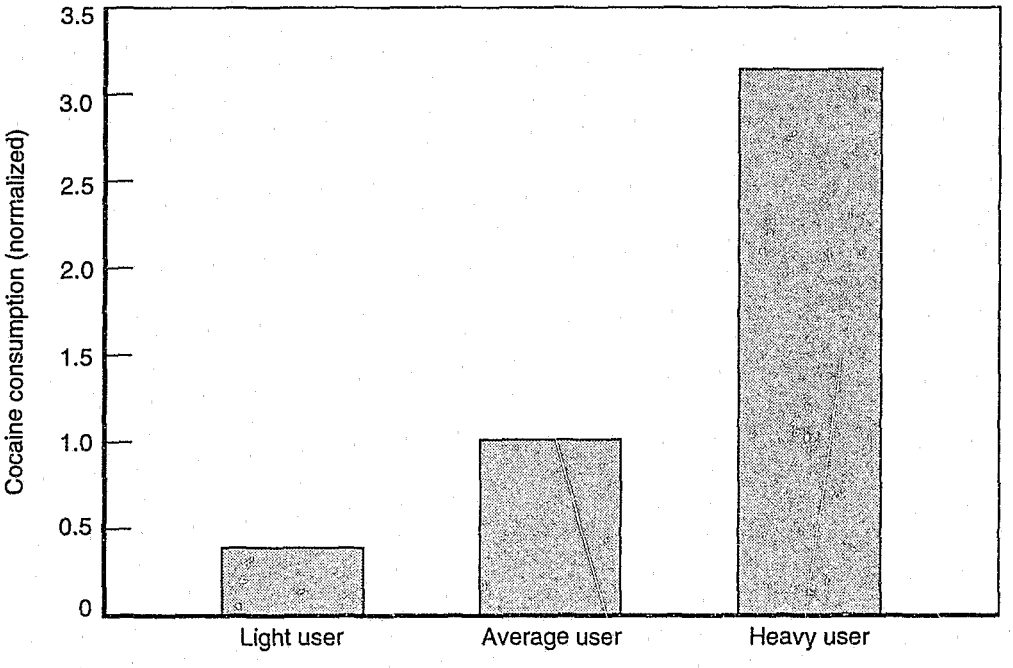

Figure 3.3—Average Annual Amount of Cocaine Consumed by Light and Heavy Users Normalized to Average Amount Consumed by Average Users

[12]Unlike the ratio of light to heavy users, this ratio of average light user consumption to average heavy user consumption is assumed constant for all years of the epidemic.

[13]See, for example, Rydell and Everingham (1994). A rough estimate of the amount of cocaine seized by law enforcement agencies is about 100 metric tons; if only 20 or so metric tons are consumed, this implies that 80 percent of all cocaine in the United States is interdicted, which is highly implausible.

[14]The number of past-month users was estimated from the NHSDA to be around 1.6 million in 1990. Since it is unlikely that this figure is off by an order of magnitude or more, it is very likely that the survey respondents' estimates of how much they have consumed in the last month are generally quite low.

As mentioned above, we assumed both heavy and light users underreport to the same degree, so the ratio of light-user to heavy-user consumption, one to eight, is justified by the above analysis, even if the actual amounts consumed on an annual basis by light and heavy users are not.

In sum, light users were defined as those who use less often than several times a month (i.e., less often than weekly), and heavy users were defined as those who use several times a month or more. The average heavy user annually consumes about eight times as much cocaine as does the average light user, and the average heavy user's consumption is more than three times the average consumption of all cocaine users. Although the consumption estimates are not relevant to establishing the Markovian model of demand, they were used to analyze consumption trends once the Markovian model was established. Exactly how much cocaine each average heavy user and each average light user consumes annually must be determined using a reasonable estimate of the total cocaine consumed in the United States in a year. This was done for the reference year 1992 (see Chapter Seven), in which consumption was estimated to be 291 metric tons (Rydell and Everingham, 1994).

VARIATION IN THE FRACTION OF HEAVY USERS OVER TIME

We assume that all new users are light users. So near the onset of the epidemic, nearly all users are light users. But with time, light users flow on to heavy use and the number of heavy users increases. There is no reason to expect that the fraction of all users that are heavy users remains constant with time, and in fact it does not. Table 3.5 reports the numbers of light and heavy users in 1985, 1988, and 1990 estimated from the corresponding NHSDA surveys, and the corresponding percentage of all users that are heavy users. The percentage of all users that are heavy users increases from 13.7 percent in 1985 to 22.0 percent in 1990. Two effects contribute to this increase: light users are flowing on to heavy use, and incidence (assumed into light use) is decreasing.

The Markovian model is required to match not only overall prevalence (the number of all users, whether light or heavy) over time (i.e., for every survey year from 1972 to 1991), but also the percentage of all users that are heavy users over time (for 1985, 1988, and 1990, the three survey years for which enough data were available to con-

Table 3.5

Percentage of All Household Users That Were Heavy Users in 1985, 1988, and 1990

	Users (millions)			
Year	Light	Heavy	Total	Percent Heavy
1985	10.3	1.6	11.9	13.7
1988	6.7	1.5	8.2	18.4
1990	4.9	1.4	6.2	22.0

SOURCES: 1985, 1988, and 1990 NHSDAs.

duct the frequency analysis[15]). Both the overall prevalence and the percentage of users that are heavy users estimated from the NHSDA were first adjusted to account for cocaine use among two populations not represented in the NHSDA, the homeless and the incarcerated. Our estimations of the number of homeless and incarcerated cocaine users are detailed in the next two sections, after which the estimates are combined to establish the overall prevalence estimates to which the model was fit.

ESTIMATES OF PREVALENCE AMONG THE HOMELESS

Presented here are very rough estimates of (1) the size of the homeless (or near-homeless[16]) population in the United States over the past several years, and (2) the number of light and heavy cocaine users within that population, which varies over time. The available estimates for the number of homeless are very broad, sometimes contradictory, and limited to just a few years; estimates of how many homeless people nationwide use cocaine do not seem to be available at all. The scarcity of good data on either the number of homeless or the prevalence of drug use among the homeless severely limited the accuracy of these estimates. We thus make no claims about them except that we believe they are reasonable and the best available.

The Number of Homeless and Near-Homeless

Estimates of the number of homeless and near-homeless people in the United States over the past three decades were derived as follows. A mid-range estimate of the number of homeless in 1983 is 300,000, reflecting Department of Housing and Urban Development (HUD), National Bureau of Economic Research (NBER), and ICF, Inc., estimates, as reported by the Urban Institute (Burt and Cohen, 1989, p. 25). The Urban Institute's estimate of the number of homeless in 1987 is approximately 500,000 based on a sampling of service-using homeless (i.e., homeless in shelters or using soup kitchens) and assuming 20 non-service-using homeless for every 100 service-using homeless.[17] Estimates for other years reported by the Urban Institute are based on nominal constant annual growth rates, but since those nominal rates are not supported by empirical evidence, we did not use those estimates.

Until 1987, a constant annual growth rate (geometric growth) was assumed, the magnitude of which was determined by the 1983 and 1987 estimates.[18] The annual growth rate based on the 1983 and 1987 estimates turned out to be about 15 percent, which is similar to those nominally assumed in other studies.[19] After 1987, linear

[15]Although the 1991 NHSDA population estimates were available for our study, the detailed data necessary to differentiate between light and heavy users by frequency of consumption were not.

[16]That is, marginally housed (see discussion below).

[17]The Urban Institute suggests that assuming as many as 50 non-service-using homeless for every 100 service-using homeless would also be reasonable. However, most studies report street-to-shelter ratios that are lower than 50/100, and service-using homeless include not only those in shelters, but also those using soup kitchens. (Burt and Cohen, 1989, pp. 29–30.)

[18]This assumption has been used in other studies, for example, Burt and Cohen (1989, p. 25).

[19]See Burt and Cohen (1989, p. 25).

growth of 50,000 per year was assumed.[20] This population includes both adults and children. The Urban Institute estimates that about 15 percent of the service-using homeless in 1987 were children;[21] assuming this fraction to be constant, the annual number of homeless adults was determined.[22]

Thus far are included the number of homeless people on the streets or in shelters at a given point during the year. It is estimated that many more experience homelessness at some time during the year—two to three or more times as many. The people who are not homeless but who have unstable housing arrangements—i.e., the *marginally housed* or *near-homeless*—are unlikely to be represented in households and are thus unlikely to be counted in the NHSDA.[23] Assuming the number of near-homeless to be 1.5 times the number of homeless,[24] and assuming the fraction of children in this population is the same as it is for the homeless, the number of homeless/near-homeless adults was calculated. The estimates of the numbers of homeless (with the two original data points indicated), homeless/near-homeless, and adult homeless/near-homeless are plotted in Figure 3.4.[25]

The Fraction of Homeless That Use Cocaine

Fischer (1987) provides some insight into the prevalence of drug use and abuse among the homeless prior to 1987. Her paper reviews a number of then-recent studies and presents the reported estimates of illicit drug use. She states (1987, p. 6):

> Since definitions and measures of drug use were not comparable in most cases, estimates were grouped in two categories consisting of reports of ever or occasionally using drugs and recent or regular use. This is a crude indicator of "casual" use versus abuse in homeless individuals. The estimates of drug use ranged from 3 percent to 31 percent.

---

[20]Estimation of the number of homeless after 1987 based on continued geometric growth led to implausibly (but not impossibly) high numbers for recent years. Because good nationwide estimates of the number of homeless in recent years were not available, we adopted the more conservative assumption of constant annual growth. Under the pre-1987 assumption of constant annual growth rate, the number of homeless increased about 0.5 million between 1986 and 1987; this was the post-1987 growth we assumed.

[21]Burt and Cohen (1989, p. 28). The authors do not explicitly define the age at which young people are considered adults. We assumed that their definition of children corresponds to people too young to use cocaine.

[22]The Institute of Medicine reports that about 75 percent of the homeless are unattached adults and the rest are mostly single mothers with children (Gerstein and Harwood, 1990, p. 84). This finding is not inconsistent with the Urban Institute estimate: 75 percent of the homeless are unattached males, 8 percent are unattached females, 8 percent are single mothers with children, 2 percent are other families with children, and the rest are other family groups without children (Burt and Cohen, 1989, p. 39).

[23]Some stay temporarily during these intervals of homelessness with family or acquaintances, but nonetheless they are excluded from the household population (Gerstein and Harwood, 1990, p. 84).

[24]The Institute of Medicine claims that "200,000 to 700,000 people . . . are homeless on any given night and as many as 2 million experience homelessness at some point during a year" (Gerstein and Harwood, 1990, p. 84). The Urban Institute calculates that about twice as many people experience homelessness at some time during the year as are homeless during a month (Burt and Cohen, 1989, p. 32).

[25]Only the counts from about 1972 and later are relevant to the analysis, since 1972 is the first NHSDA survey year for which cocaine data are available. We included the earlier years so that our model of the cocaine epidemic would have an initial tail instead of an abrupt start.

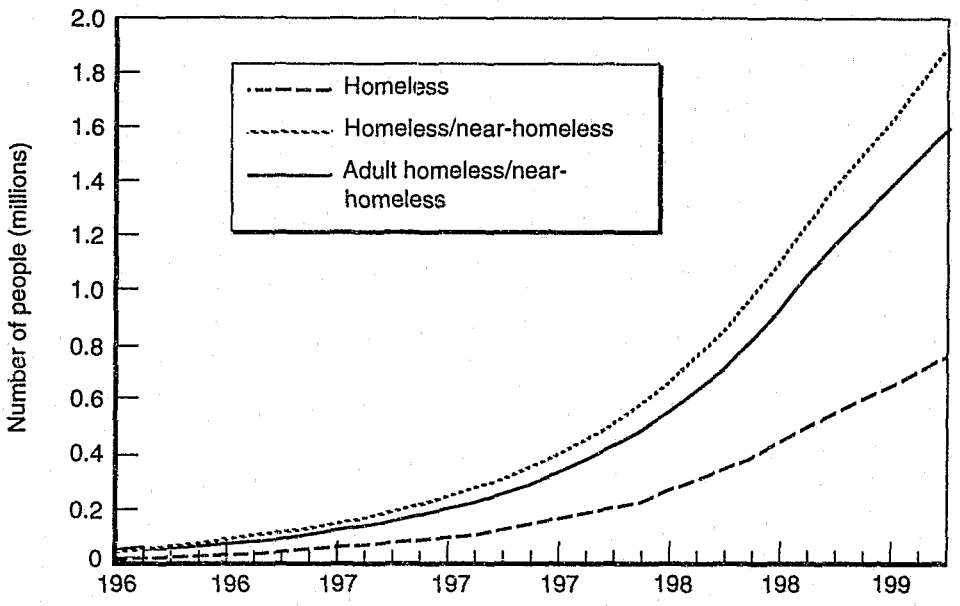

Figure 3.4—Numbers of Homeless, Homeless and Near-Homeless, and Adult Homeless/Near-Homeless Over the Past Two Decades

For lack of better information, we assumed that ever/occasional use corresponded to light use, and recent/regular use corresponded to heavy use. Only one of the studies covered by Fischer (one published in 1987 and thus presumably representing 1986 conditions) was national. It reported (Fischer, 1987, Table 5) that 10 percent of the nation's homeless were recent/regular drug users, but did not report what percentage were ever/occasional users. Averaging the ratio of ever/occasional to recent/regular percentages in those studies that did report both (each of which focused upon a particular city) suggests that it is reasonable to estimate that the percentage of ever/occasional drug users among the nation's homeless was around 20 percent (twice the recent/regular prevalence rate) in 1986.

How much of that drug use can be attributed to cocaine? Fischer states (1987, p. 2) that "although alcohol is the drug of choice among the homeless, partly due to economics, there is evidence suggesting that [illicit] drug abuse also affects substantial proportions." Prior to the introduction of crack, which was sometime before 1987 (the first year the NHSDA surveyed crack usage), cocaine was probably not widely used by the homeless. Assuming that one-fourth of the illicit drug use among the homeless in 1986 can be attributed to cocaine, we estimate that about 5 percent of the homeless were light cocaine users and 2.5 percent were heavy cocaine users in that year. These prevalence rates are comparable to (although a bit higher than) the estimates (based on the NHSDA) of the prevalence of light and heavy cocaine users in the household population. Prevalence rates among the homeless for years prior to 1986 were determined by adjusting the 1986 light and heavy prevalence rates (5 per-

cent and 2.5 percent, respectively) by rough estimates of the light and heavy prevalence rates in the household population.[26]

Estimating the extent of cocaine use among the homeless for the years after 1986 is more difficult. According to unpublished data collected by Audrey Burnam of RAND, there is evidence that a dramatic increase in drug usage and dependence among the homeless in Los Angeles occurred recently (from 1985 to 1991), presumably *because of an increase in cocaine, and particularly crack, usage.* If this pattern is also true nationally, then assuming that the cocaine usage rates after 1986 are the same as in 1986 leads to a serious underestimate of cocaine usage among the homeless. Burnam's data show that recent drug use (within the past six months) among the homeless in Los Angeles increased from 10 percent to 29 percent, lifetime use increased from 31 percent to 51 percent, and a startling 21 percent of the homeless in Los Angeles are *dependent* on cocaine. For our analysis, it was assumed that both the light and heavy prevalence rates increased linearly to 20 percent in 1991. (If drug usage among the homeless has increased because of crack addiction, assuming the light rate is twice the heavy rate is no longer justifiable.)

The Numbers of Light and Heavy Cocaine Users in the Homeless/Near-Homeless Population

Finally, we assumed that the prevalence rates among the near-homeless are one-half the prevalence rates among the homeless.[27] Figure 3.5 depicts the numbers of light and heavy cocaine users in the homeless/near-homeless population for each year of the cocaine epidemic. This estimation suggests that the numbers of light and heavy cocaine users were not significant prior to 1986, but that they became increasingly significant after 1986. Hereafter, the combined homeless/near-homeless population will be referred to as simply the homeless population.

ESTIMATES OF PREVALENCE AMONG THE INCARCERATED

Estimation of the number of cocaine users among those who are incarcerated is also a two-step process. First, the size of the incarcerated population for each year in the past three decades is assembled, and then the fractions of the incarcerated population that are light and heavy cocaine users, which vary over time, are roughly estimated. These figures determine the number of light and heavy cocaine users among the incarcerated that must be added to the NHSDA-derived counts. By incarcerated cocaine users, we mean people who would be users if they were not incarcerated. Incarcerated people consume little if any cocaine, reflecting the incapacitation effect of incarceration.[28] However, those people using drugs before entering jail or prison

[26]The rough estimates were derived from an early version of the Markovian model (fit to NHSDA data only, and unadjusted for homeless and incarcerated users). Using the household prevalence rates, instead of the prevalence numbers, ensures that homeless prevalence rates during the early years of the cocaine epidemic are not overestimated.

[27]The Institute of Medicine report (Gerstein and Harwood, 1990, p. 84) makes a similar assumption.

[28]Between July 1, 1989, and June 30, 1990, 0.4 percent of drug tests in federal prisons and 1.4 percent of drug tests in state confinement facilities were positive for cocaine. However, these numbers may

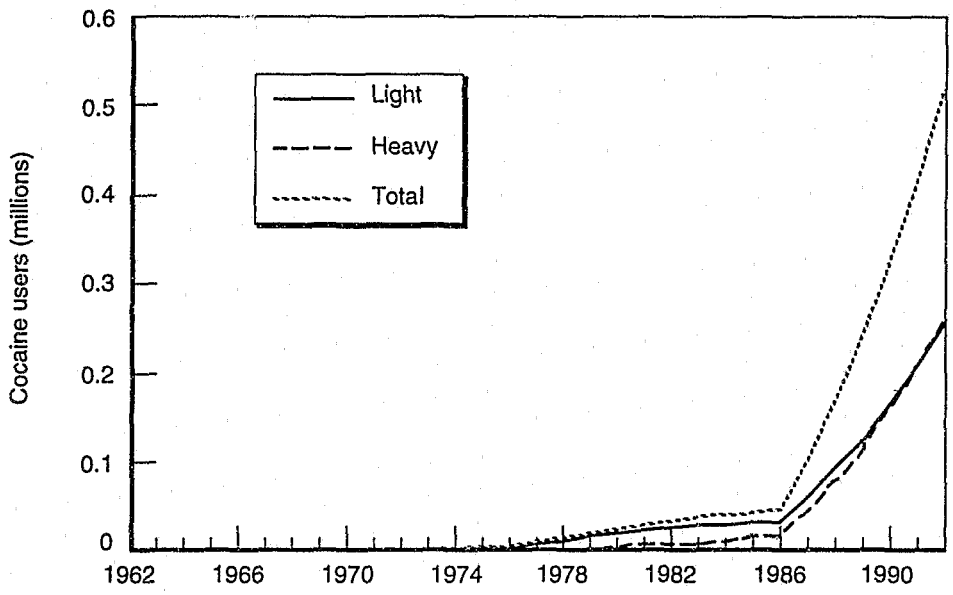

Figure 3.5—Light, Heavy, and Total Cocaine Users in the U.S. Homeless/Near-Homeless Population

will likely use drugs after release, unless treated for drug abuse/addiction while incarcerated. We thus considered people who were drug users before incarceration to be drug users, even though they did not use drugs while incarcerated.

The Size of the U.S. Incarcerated Population

The numbers of people in federal prison, state prison, or jail for each year from 1960 to 1990 are displayed in Table 3.6. The numbers represent average prisoner counts on any given day during the year, not the total number of people cycled through the system in a given year. As such, they represent the size of the population not counted in the NHSDA surveys. These data (combined with estimates of the missing data) are then plotted in Figure 3.6.

The Fraction of the Incarcerated Population That Use Cocaine

An analysis of the data from the 1986 Survey of Inmates of State Correctional Facilities showed that 43.7 percent of state prison and jail inmates admitted to having ever used cocaine, and that 22.2 percent admitted to having used cocaine regularly (once a week or more). Unfortunately, these data do not correlate directly with the NHSDA-derived data, since the inmates were not asked if they had used cocaine in

somewhat overstate the actual prevalence because they include tests that were for cause, not just random and systematic screens (Bureau of Justice Statistics, 1991).

Table 3.6

Counts of Various Incarcerated Populations (in millions)

Year	Federal and State Prison	Federal Prison	State Prison	State Jail	Total[a]
1960	0.212	0.023	0.189	0.119	0.331
1961	0.220				0.342
1962	0.218				0.344
1963	0.217				0.346
1964	0.214				0.346
1965	0.210	0.021	0.189		0.346
1966	0.199				0.338
1967	0.194				0.336
1968	0.187				0.332
1969	0.196				0.345
1970	0.196	0.020	0.176	0.152	0.348
1971	0.198				0.345
1972	0.196			0.141	0.337
1973	0.204				0.348
1974	0.218				0.365
1975	0.240	0.024	0.216		0.390
1976	0.262				0.414
1977	0.278				0.433
1978	0.294			0.158	0.452
1979	0.301				0.472
1980	0.315	0.020	0.295		0.499
1981	0.353				0.550
1982	0.394	0.023	0.371		0.604
1983	0.419	0.026	0.393	0.223	0.642
1984	0.443	0.027	0.415	0.234	0.676
1985	0.480	0.032	0.447	0.256	0.735
1986	0.522	0.036	0.485	0.274	0.795
1987	0.560	0.039	0.521	0.295	0.855
1988	0.603	0.042	0.560	0.343	0.945
1989	0.680	0.047	0.633	0.395	1.075
1990	—	0.050	0.688	0.405	1.143

SOURCES: (1) Bureau of Justice Statistics, *Source Book of Criminal Justice Statistics, 1990*, p. 604; (2) *Statistical Abstracts of the United States*, 1967, p. 163; 1972, p. 162; 1975, pp. 167–168; 1992, p. 197.

[a]Summation of federal and state prison and state jail populations (columns 2 and 5). The number of people in state jails for years without data is estimated by linear interpolation of the state jail data.

the past year (data we used to determine prevalence), and the question about regular use did not specify how recently that regular use had occurred. However, since past-year users are a subset of lifetime users and recent weekly users are a subset of people who have used weekly at some point, 43.7 percent and 22.2 percent are upper bounds on the fractions of inmates (in 1986) that were users and heavy users of cocaine, respectively. The fraction of all inmates that are light or heavy cocaine users was estimated using these upper bounds.

For years after 1986, we assumed the fraction of all inmates that are light users (21.5 percent) and the fraction that are heavy users (22.2 percent) remained constant. For

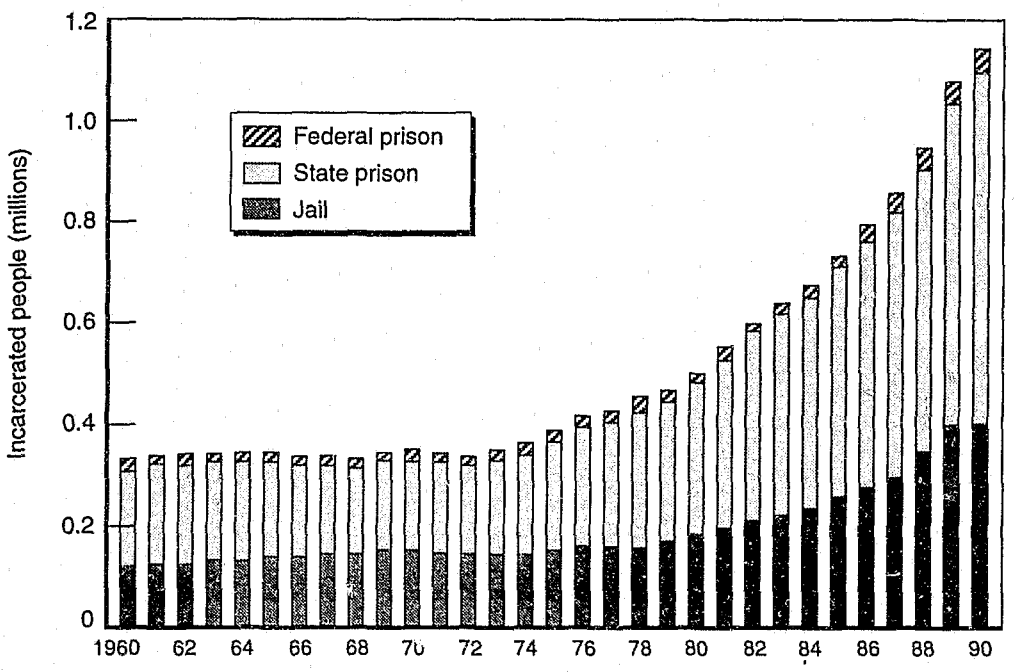

Figure 3.6—Federal Prison, State Prison, and Jail Populations in the United States over Time

years prior to 1986, these fractions were adjusted to be proportional to estimated light and heavy prevalence numbers.[29]

The Numbers of Light and Heavy Cocaine Users in the Incarcerated Population

Combining the size of the incarcerated population with the fractions representing light and heavy cocaine users determines the number of light and heavy cocaine users among the incarcerated. Figure 3.7 shows the numbers of light users and heavy users and the total for each year of the recent cocaine epidemic. The numbers become gradually more significant with time, reflecting in part the rapid increase in the prison population since the late 1970s. The numbers for years after 1990 are assumed to be the same as for 1990.

OVERALL PREVALENCE ESTIMATES

The homeless estimates and prison/jail estimates were combined with NHSDA-derived estimates of light and heavy cocaine users to derive total prevalence esti-

[29]As was done for the homeless estimation, the estimates were derived from an early version of the Markovian model (fit to NHSDA data only, and unadjusted for homeless and incarcerated users).

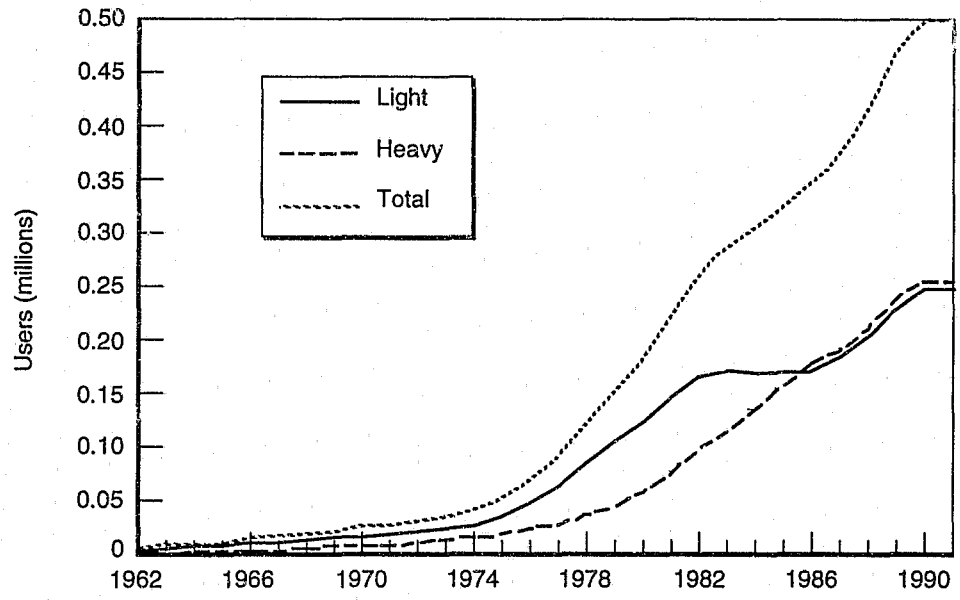

Figure 3.7—Light, Heavy, and Total Cocaine Users in the U.S. Incarcerated Population

mates.[30] The overall (light and heavy together) prevalence for NHSDA survey years is shown in Figure 3.8; the Markovian model was fit to these combined data.

The fraction of cocaine users that are heavy users was determined above using only NHSDA data for 1985, 1988, and 1990. A greater proportion of both the homeless and the incarcerated cocaine users are heavy users in our estimation. Therefore, the NHSDA-derived fractions were adjusted to account for these two additional populations. The fractions of all three populations were weighted by the population sizes and averaged in order to determine the fractions to which the Markovian model was fit. These adjusted fractions are displayed in Table 3.7.

It should be noted that some nonhousehold populations remain excluded or not fully included: (1) the institutionalized and hospitalized, which is a very small population; (2) military personnel living in military quarters, a population that presumably exhibits a low prevalence of drug use by virtue of its regimented lifestyle and pervasive drug testing; and (3) college dormitory residents, of which there are over 2 million. (The 1991 NHSDA was the first to survey some nonhousehold populations, including college dormitory residents and the sheltered homeless, but it did not survey the military, the institutionalized, or the transients.) Although the estimates of the num-

[30]Another minor adjustment to the 1991 prevalence numbers was needed to reflect the fact that the 1991 NHSDA for the first time surveyed the homeless in shelters and to avoid double counting this group. Of the estimated 248,000 homeless and near-homeless light cocaine users, 142,000 are homeless (and not near-homeless). Of these 142,000, 118,000 (5/6) use services (consistent with the assumption that there are 20 non-service-using homeless for every 100 service-using homeless). The Urban Institute (Burt and Cohen, 1989, p. 38) estimates that 3/4 of the service-using homeless use shelters. Thus, 89,000 (3/4 of the 118,000 service-using homeless light cocaine users) use shelters. Thus, the total number of light users in 1991 must be reduced by 89,000. The reduction in the number of heavy cocaine users is the same.

ber of drug users in the United States might improve somewhat if the counts (for years prior to 1991) were adjusted to reflect the college dormitory resident population, this adjustment, if possible at all, would be at best a rough guess, and certainly is not critical for cocaine.[31]

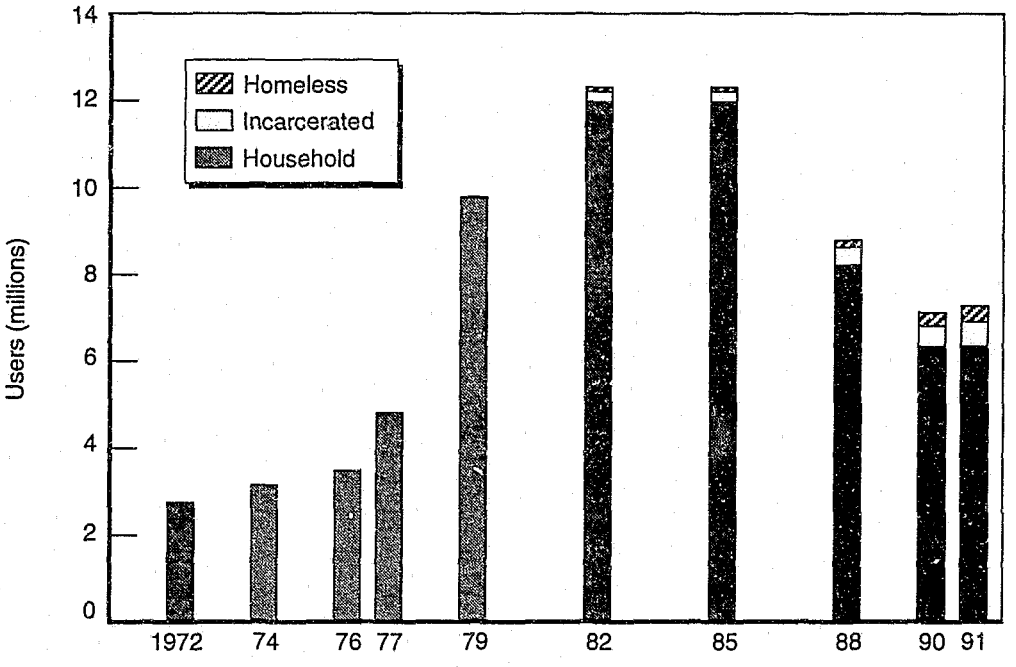

Figure 3.8—Overall Prevalence of Cocaine Users in the United States

Table 3.7

Percentage of All Users That Were Heavy Users in 1985, 1988, and 1990

Year	Household[a]	Homeless[b]	Incarcerated[b]	Total
1985	13.7	30.8	47.9	14.6
1988	18.4	46.3	50.8	20.4
1990	22.0	49.3	50.8	25.3

[a]From 1985, 1988, and 1990 NHSDAs.

[b]Derived from the estimates above.

[31]However, it might be more critical for alcohol and marijuana. As a result of including dormitory residents, the NHSDA prevalence rates of cocaine users aged 18 to 25 are only slightly lower, but the prevalence rates of alcohol and marijuana users are higher (National Institute on Drug Abuse, 1991).

Chapter Four

COCAINE USER COHORT RETENTION

There is a well-understood fact about illicit drug use: many users, as they mature, naturally desist in using drugs, and only some users continue to use drugs for a long period of time. One may ask, Of a cohort of people who all begin using drugs at approximately the same time, how many will still be using drugs one, two, five, ten, etc., years later? The fraction of a cohort of initiates that is still using drugs N year later is called the N-year cohort retention rate. This is another characteristic of the system (in addition to prevalence) that can be used to describe the dynamics to be modeled.[1]

Cohort retention rates can be calculated from the NHSDA in the following way. For each person who responded positively to the lifetime cocaine-use question, subtracting the person's age at first use from his current age determines the number of years since initiation. This establishes a set of cohorts of people who initiated use at the same time. The fraction of people in each cohort that are still using cocaine is determined by examining the responses to the past-year use question. This procedure looks at a cross section of the population for each year since initiation, combining people of different ages, races, incomes, sexes, etc. Thus, the N-year retention rate cannot necessarily be interpreted as the likelihood that an individual user will continue to use for N years, since different subgroups probably exhibit different retention rates. It can be interpreted as an average characteristic of the drug-consuming population.

As attitudes about drug use change, so do retention rates. Therefore, we might expect that cohort retention rates calculated with NHSDA data from different survey years will vary. Each curve in Figure 4.1 plots cohort retention rate as a function of time (smoothed with a three-year running average).[2] Three of the curves are derived from different years of the NHSDA; one plots the average of those three years. The data for cohorts that initiated use over 15 years ago are too noisy to be useful. The retention rates seem to have declined between 1985 and 1990, but the estimates are imprecise, so this observation is made with some caution.

---

[1] It is logically equivalent to the reciprocal of the time required for (nearly) all users to flow out of use in the absence of incidence.

[2] By definition, 100 percent of a cohort is still using zero years after initiation. Smoothing the data distorts the zero-year data point.

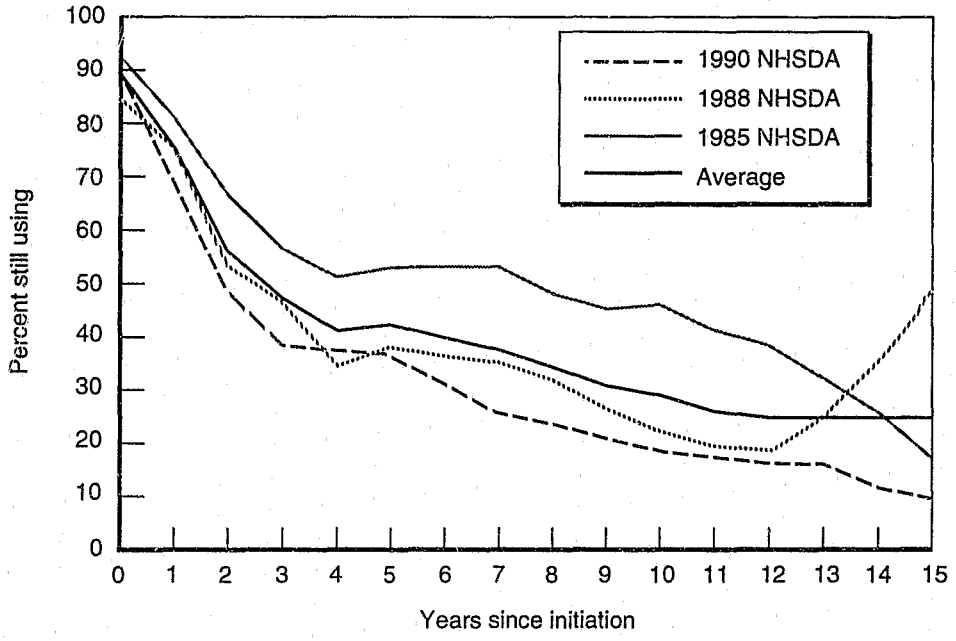

Figure 4.1—Cohort Retention Calculated from Three NHSDAs and the Average

Assuming that the average cohort retention rate curve best characterizes the recent cocaine epidemic, we used this averaged information to fit the Markovian model. By examining the average curve, we see that about 50 percent of initiates are still using two years later, about 40 percent are still using five years later, and about 30 percent are still using ten years later. In other words, a large fraction of the users only use for a short time, but those who continue to use do so for many years. The former category correlates with the experimental users, whereas the latter corresponds to the habitual and addicted users.

Chapter Five

ESTIMATES OF THE INCIDENCE OF COCAINE USE OVER TIME

Incidence of cocaine usage has varied greatly during the past 20 years. It is hypothesized that illicit drug use is "transmitted" to non-users (much like an infectious disease, hence the phrase "cocaine epidemic") by drug consumers in the early stages of use. These users, who have not yet experienced the most objectionable consequences of drug consumption, proselytize their friends with descriptions of euphoria and protestations of drug usage's social acceptability. However, since the exact nature of the transmission mechanism is not fully understood, predicting incidence of drug use is much more complicated than predicting incidence of other infectious pathologies. Even measuring incidence is difficult, given the illicit nature of drug consumption; it is not surprising that there is no useful direct count of annual incidence.[1]

Nonetheless, incidence is a critical component of the system to be modeled. It will be shown that a pronounced peak in incidence (of this most recent epidemic) preceded by almost a decade the peak in prevalence that occurred in the mid-1980s. In fact, prevalence is so closely tied to incidence (although with an inherent time delay) that an assumption of constant incidence would preclude a meaningful match of the model to the dynamics of the cocaine epidemic. Thus, to model the dynamics of the epidemic requires detailed incidence information over time. Although direct counts of annual incidence for the entire duration of the epidemic are not available, annual incidence can be derived from the NHSDA in either of two straightforward ways.

The NHSDA asks subjects if they have used cocaine in the past 30 days, in the past year, or ever in their lifetime. The difference in the lifetime estimates between successive surveys represents incidence between surveys. The surveys until recently were administered intermittently rather than annually, so annual incidence was determined by dividing the between-survey incidence by the number of intervening years.[2] These data were smoothed using a three-point moving average to generate

[1]The 1974–1982 NHSDAs included direct questions on subjects' first-time use of drugs during the past year, but since these questions were not included in the more recent surveys, this method was not used in more than a comparative sense for our analysis.

[2]Adjustments to account for the specific months in which successive surveys were administered were not incorporated in this analysis. As discussed in Chapter Three, the fact that the lifetime prevalence estimate in 1988 was lower than in 1985 (implying negative incidence!) was corrected by interpolating the lifetime prevalence data so that incidence could be derived for all years. (See Gfroerer and Brodsky, 1992, for an alternative difference estimation.)

the *difference estimate* of the annual incidence for each year between the first and the most recent survey. (The number of users who responded positively to the lifetime-use question the first time the survey was administered in 1972 was assumed to reflect a constant annual incidence between the nominal start of the epidemic in 1962 and the time of the 1972 survey.[3])

Those who respond positively to the lifetime-use question are asked at what age they began to consume cocaine. This information, the respondent's age at the time of the survey, and the date of the survey can be used to determine the year of first use for each respondent, which can then be compiled over several survey years (1985, 1988, 1990, and 1991) to generate the *retrospective estimate* of annual incidence for each year of the cocaine epidemic from 1962 through 1989 (Gfroerer and Brodsky, 1992). The trend from shortly before 1989 was linearly extrapolated to estimate more recent incidence.

Each method is subject to error. The main advantage of the retrospective method is that it is based on a larger sample size, which tends to stabilize the estimate. A third estimate, the *average estimate,* was determined by simply averaging the difference and retrospective estimates. Averaging the two estimates mitigated the potential errors of the separate estimates.

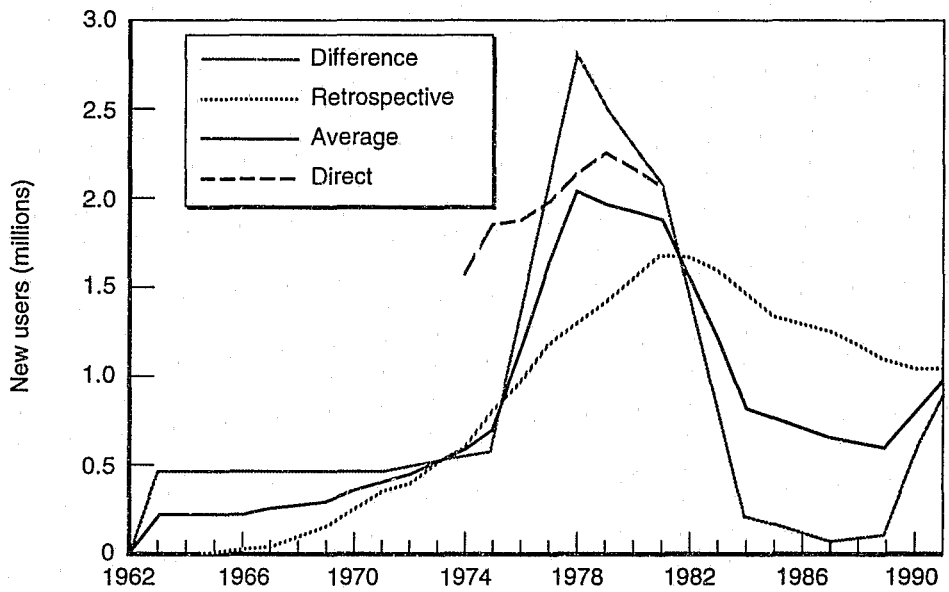

Figure 5.1—Estimation of Annual Incidence: Comparison of Four Methods

[3]Neither the year of the start of the epidemic nor the shape of the incidence curve before the first survey year is critical, but assuming a gradual smooth start to the model of the epidemic avoids the artificial boundary effect that would result from simply assuming all users started right before the first survey.

Annual incidence of cocaine usage as determined by these three methods is plotted in Figure 5.1. Also plotted is the *direct estimate* for the years between 1974 and 1982 for the purposes of comparison.[4] There are significant differences between the various estimates; for example, the incidence peaks are displaced by as much as five years, and the retrospective estimate displays no evidence of a recent upturn in the incidence of cocaine usage.

Initial results from this demand modeling analysis indicated that neither the retrospective nor the difference estimate of incidence provided acceptable model parameter estimation. Thus, because it is intermediate to the other two estimates, the average estimate was used.

---

[4]These data were extracted from Gfroerer and Brodsky (1992). Data for missing years between 1974 and 1982 were determined by linear interpolation.

FITTING THE MODEL

The model to which the observed data were fit is a two-state, four-parameter Markovian model (see discussion in Chapter Two and depiction in Figure 2.2). This model was chosen because it is the simplest sufficiently detailed model capable of generating the requisite historical trends.

THE FITTING PROCEDURE

The nature of the observed data—noisy, imprecise, and sparse—precludes the effective employment of a rigorous fitting procedure (such as a regression). Instead, the four-dimensional parameter space[1] was exhaustively searched for choices that best matched data identified as characterizing the system, the definition of "best match" being, admittedly, somewhat subjective (see below).

Three types of information about drug usage were utilized in the parameter estimation procedure: total prevalence of cocaine use (light and heavy together) over time, fraction of all cocaine users that are heavy users, and cohort retention rate. (The first two of these were defined and the observed data described in Chapter Three; the third was discussed in Chapter Four.)

- *Total prevalence over the course of the epidemic.* Specifically, the prevalence estimates from the ten survey years from 1972 to 1991 were compared to the modeled prevalence estimates from those same ten years. The mean squared error between the observed and the modeled prevalence was the measure of merit.
- *Fraction of heavy users over recent time.* The proportion of all cocaine users that are heavy users is not constant because all new users are light users, and new heavy users originate only from the light-user pool. The fraction of heavy users over three recent survey years, 1985, 1988, and 1990, increased from 0.15 to 0.25. The modeled fraction of heavy users was compared to the observed fractions for those three years.

[1] By definition, a, b, f, and g must be between 0.0 and 1.0. Constraints in the model reduce the size of the parameter space even further; for example, $a + b$ and $f + g$ must both be less than 1.0.

- *Ten-year cohort retention rate.* As discussed in Chapter Four, cohort retention rate is the fraction of a cohort of initiates that will still be using cocaine after some period of elapsed time. Cohort retention rates can be determined from the NHSDA utilizing the age-at-first-use data of the lifetime users and their responses to the question about use in the past year. The ten-year cohort retention rate of the model was required to match the average of the observed 8-, 9-, 10-, 11-, and 12-year average (i.e., averaged over three survey years) cohort retention rates, which was close to 29 percent.

There is no obvious way to define a single measure of the goodness-of-fit of the model by combining these data. Since a perfect fit of all the data is generally quite unlikely (and in this case was discovered to be impossible), criteria for defining the best possible fit were needed. The fitting procedure required that the mean square error over the ten survey years be near-minimal. (The model parameters that correspond to the minimal mean square error were close to, but not the same as, those that optimized the other two measures of merit.) The fitting procedure also required the model to reproduce the trend and to approximately match the three fractions of heavy users (for 1985, 1988, and 1990). The ten-year cohort retention rate was required to be as close to the observed value as the discretization of the four-dimensional parameter space supported.

To search the four-dimensional parameter space, first f and g were fixed and a and b were varied (with step sizes of 0.005 and 0.002, respectively). The best a and b for the fixed f and g were selected. Then f and g were varied (with step sizes of 0.01 each), and the process of selecting the best a and b was repeated. Finally, the overall best set of parameters a, b, f, and g was selected.

More important than the details of the fitting procedure is a demonstration that the selected parameters lead to a good fit of the model to the observed data, and an illustration of the sensitivity of the fit to variation in the parameters.[2] Figure 6.1 shows the sum squared prevalence delta (which is proportional to the mean squared delta, or error) versus both a and b for fixed f and g (the fixed values are those that ultimately were selected).[3] The elevation plot shows that the sum squared delta must be greater than 10.0 and that it is minimized for values of a and b corresponding to the middle band.

Figure 6.2 plots the ten-year cohort retention rate as a function of a and b for the same fixed f and g. The diagonal line at the bottom of the darkest band in the elevation plot corresponds to a retention rate of 29 percent. Crossing this plot with the previous elevation plot (bottom, Figure 6.1) determines a set of values for a and b that are pretty good (for these fixed f and g).

---

[2]In addition, it is comforting to see that the functions are very well behaved (i.e., not at all erratic).

[3]The three-dimensional plot (the top one) allows visualization of the surface, whereas the accompanying elevation plot (the bottom one) allows easier determination of the functional values.

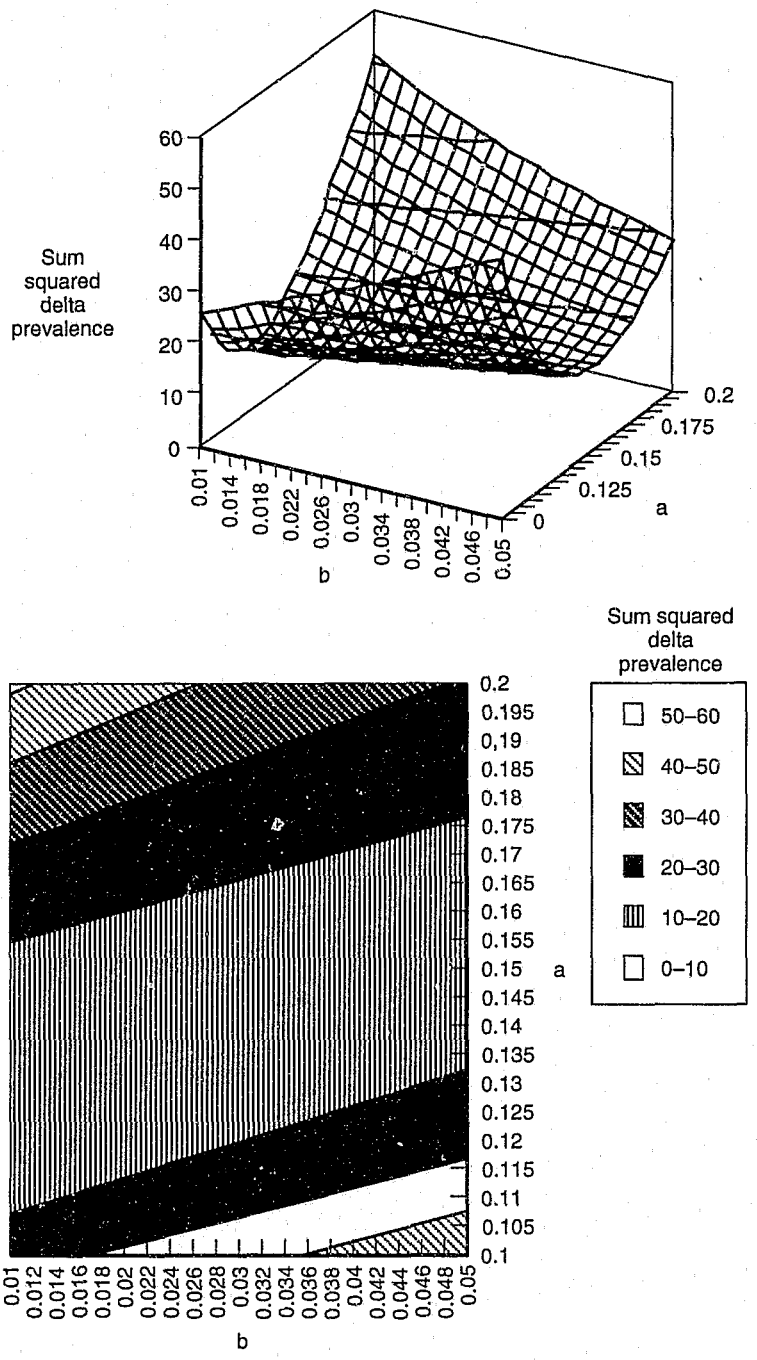

Figure 6.1—Sum Squared Delta Prevalence for Fixed $f = 0.04$ and $g = 0.02$

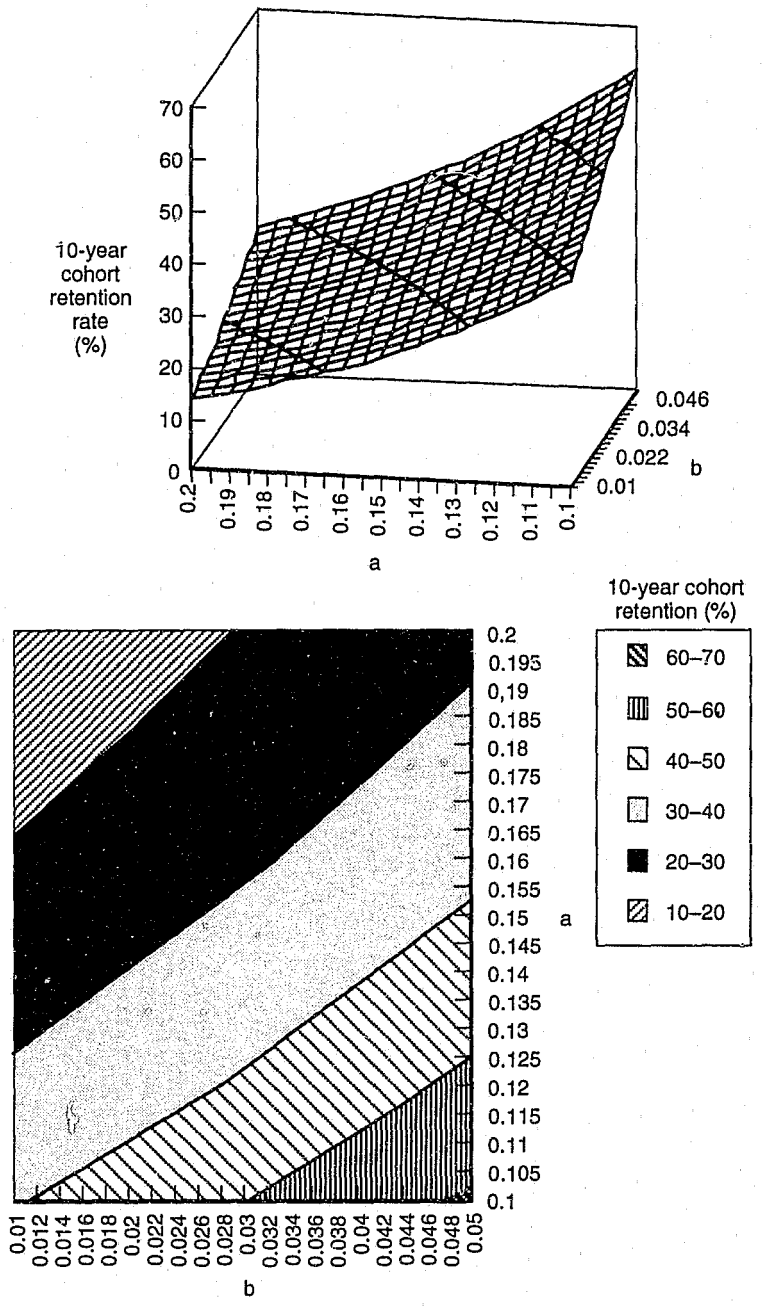

Figure 6.2—Ten-Year Cohort Retention for Fixed $f = 0.04$ and $g = 0.02$

The search reveals that for all near-optimal f and g, the value of a must be about 0.15. In Figure 6.3, the modeled fraction of all users that are heavy users as a function of b is plotted for fixed a, f, and g. It is apparent that no single value of b satisfies the requirement of matching the observed values for all three years. Any value between the vertical arrows is an acceptable compromise.

This analysis was then repeated, this time for fixed a and b, varying f and g. Figure 6.4 illustrates that to minimize the sum (or mean) squared delta in the prevalence estimate, the value of g must be quite small.

Figure 6.5 plots the ten-year cohort retention rate as a function of f and g for the same fixed a and b. In the elevation plot (bottom), the line at the top of the darkest band (the fourth band from the lower left corner of the plot) corresponds to a 29 percent ten-year cohort retention rate. As before, crossing the two elevation plots (in Figures 6.4 and 6.5) leads to a set of pretty good values for f and g (for these fixed values of a and b, 0.15 and 0.024, which were the values ultimately chosen).

Figures 6.6 and 6.7 plot the fraction of all users that are heavy users for three fixed parameters (a, b, and either f or g) and as a function of the fourth (g or f). Note that as either f (in Figure 6.6) or g (in Figure 6.7) increases, the curves begin to merge. This suggests that low values of both f and g are necessary for the model to adequately reproduce the observed trend in the fraction of all users that are heavy users.

Figures 6.1 through 6.7 illustrate the fitting procedure and demonstrate the adequacy of the values of the parameters chosen for the model: $a = 0.15$, $b = 0.024$, $f = 0.04$, and $g = 0.02$. Moreover, these figures roughly illustrate the sensitivity of the model to

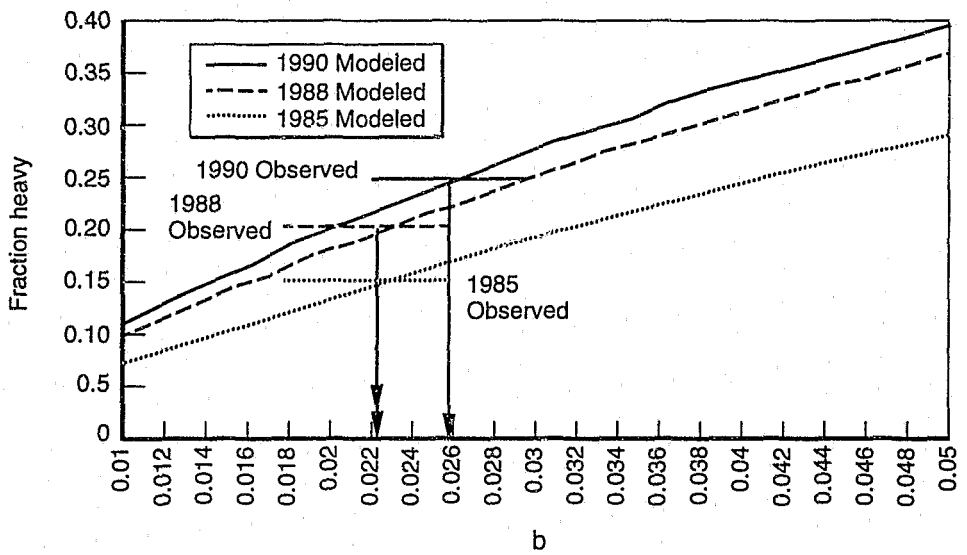

Figure 6.3—Fraction of Users That Are Heavy Users for Fixed $f = 0.04$, $g = 0.02$, and $a = 0.15$

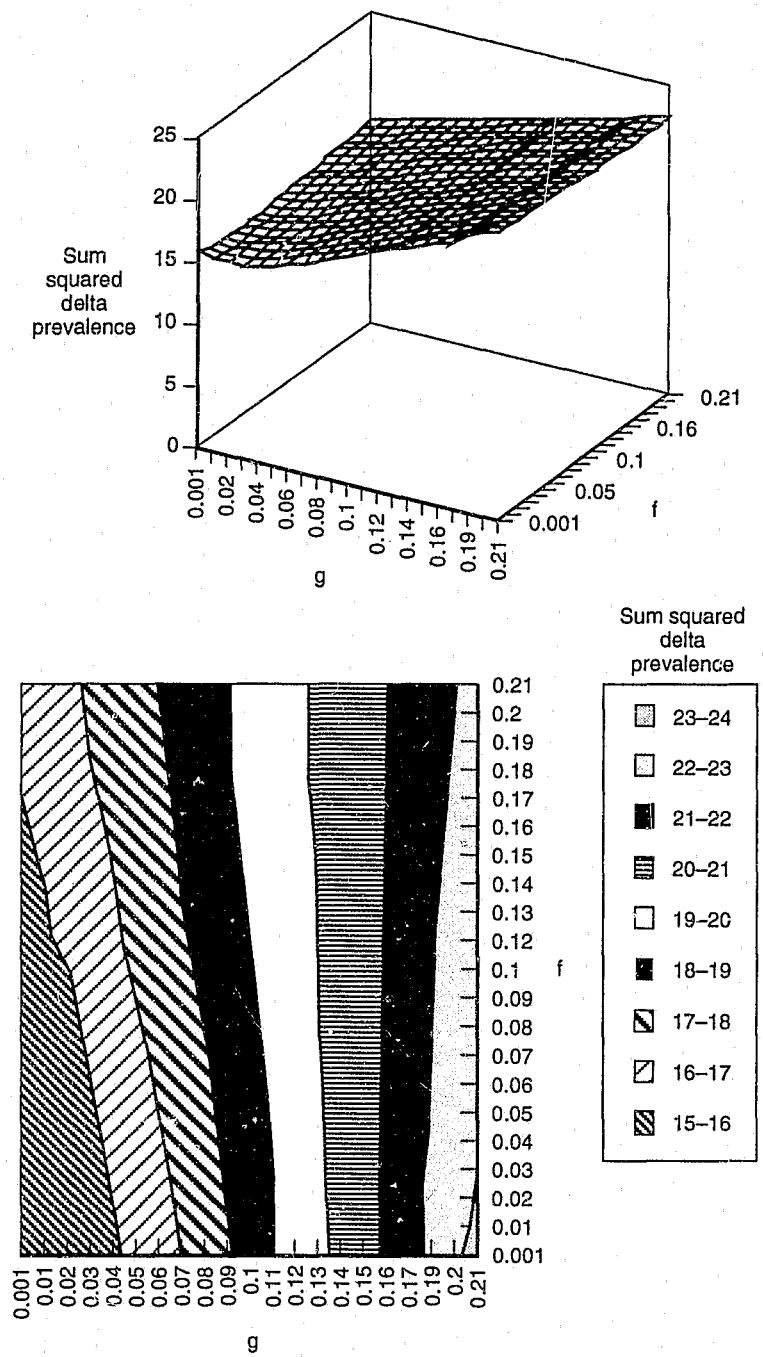

Figure 6.4—Sum Squared Delta Prevalence for Fixed $a = 0.15$ and $b = 0.024$

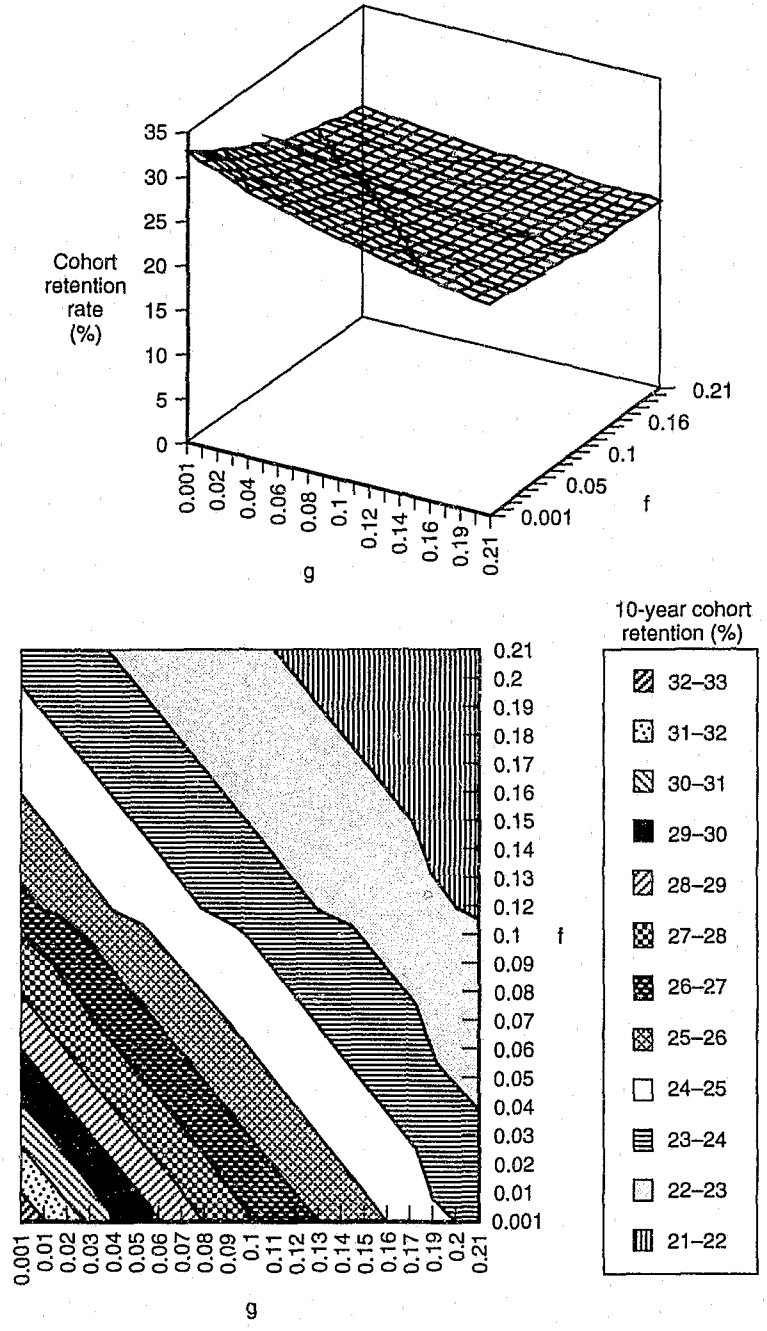

Figure 6.5—Ten-Year Cohort Retention for Fixed $a = 0.15$ and $b = 0.024$

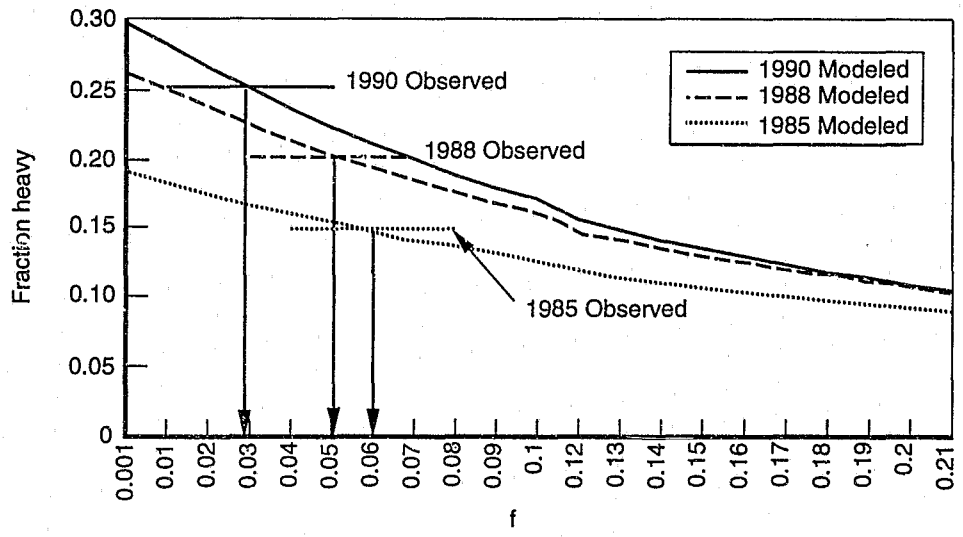

Figure 6.6—Fraction of Users That Are Heavy Users for Fixed $a = 0.15$, $b = 0.024$, and $g = 0.02$

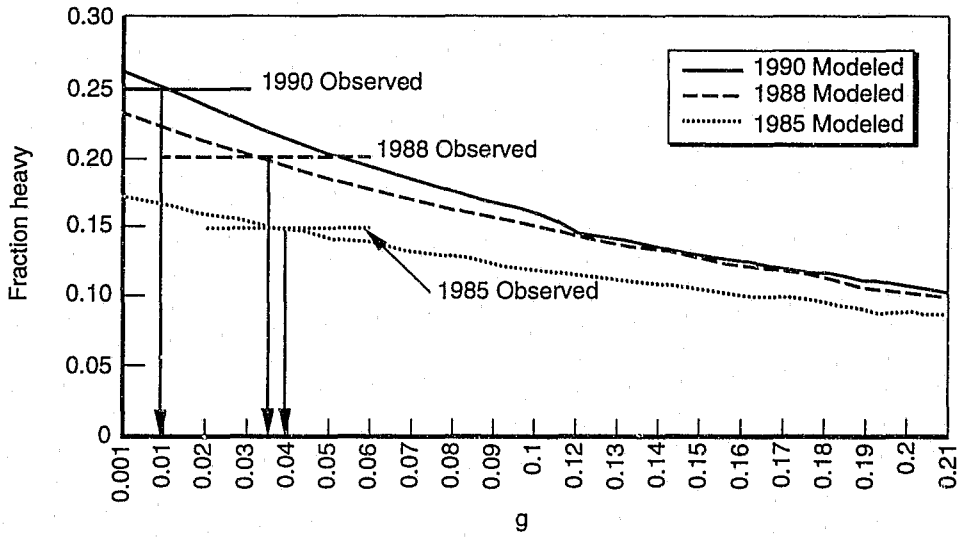

Figure 6.7—Fraction of Users That Are Heavy Users for Fixed $a = 0.15$, $b = 0.024$, and $f = 0.04$

variation in the parameter values. The set of parameter values chosen for the model is not uniquely determined, since many other combinations of values for the four parameters also provide equally good fits to the data. The ranges of values that lead to equally good fits are quite limited, however. Any choices of f and g such that the sum of the two parameters is less than 0.09 are adequate; the choices of f and g that best match the increasing trend in the percentage of users that are heavy users are those whose sum is less than 0.06. For all adequate choices of f and g, a must be about 0.15 and b must be somewhere between 0.02 and 0.03.

The results of the overall fitting procedure suggest that the outflow from heavy use to non-use (represented by the parameter *g)* is not critical to the model. A simpler model with only the backflow from heavy use to light use (represented by the parameter *f)* and the two flows out of light use (*a* and *b)* would be sufficient to fit the data.

THE FITTED MODEL

As summarized in Figure 6.8, the parameter values that make the model best fit the data are $a = 0.15$, $b = 0.024$, $f = 0.04$, and $g = 0.02$.[4] A combined outflow from heavy cocaine use of 6 percent per year seems low because it implies an average heavy-use career of about 17 years. However, the estimated parameter values are those that enable the model to best replicate the observed data (see Figures 6.1 through 6.4). Below we compare the estimates generated by this model with the observed data.

Figure 6.9 plots modeled and observed overall prevalence (light and heavy prevalence together). The fit to the overall prevalence data is about as good as could be expected, and is sufficient in light of the uncertainty surrounding the prevalence estimates. The overall shape of the prevalence curve is correct, the time of peak prevalence suggested by the model is close to that indicated by the data, and the modeled average prevalence (6.7 million) is quite close to the observed average prevalence (7.2 million).[5] Not surprisingly, the model cannot exactly reproduce the prevalence data. Because the model is smooth by design, it tends to lower the peaks and raise the valleys in the prevalence curve, as indicated in Figure 6.9.

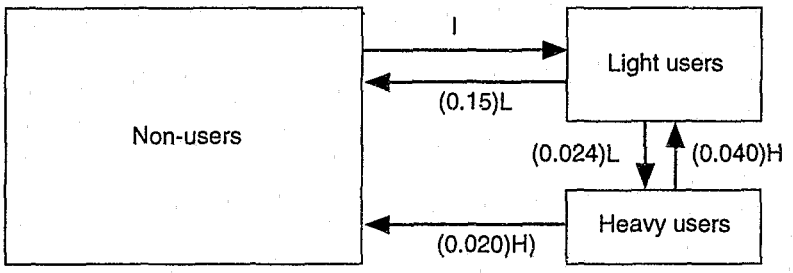

Figure 6.8—The Fitted Model

[4]Although these flow rates are traditionally called transition probabilities, they must be interpreted with great caution. For example, *a* is the annual flow rate out of light use. In a sense, it is the overall probability that a light user will flow back to non-use in any given year. However, the assumption that all initiates have a 0.15 probability of quitting each year is a misinterpretation, in part because the light-user pool also contains the backflow from heavy use. It is important to note that the parameters were chosen without regard to any data about the likely behavior of cocaine users; they were selected solely based on the fit of the two-state Markovian model to the three criteria discussed above. That the fitted model is a very simplified model of the system suggests that the four parameters are just that, parameters, possibly without deeper meaning.

[5]The average is over the ten survey years.

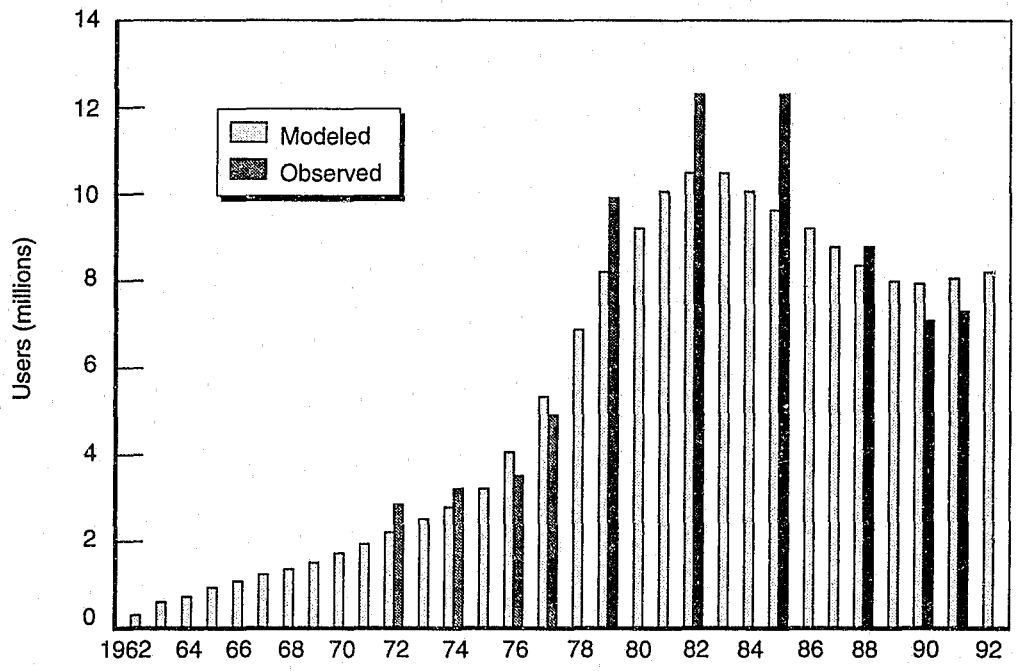

Figure 6.9—Overall Prevalence of Cocaine Users: Modeled vs. Observed

Figure 6.10 shows that the observed percentage of users that are heavy users increased from 15 percent to 25 percent from 1985 to 1990, and that the modeled percentage increased from 16 percent to 24 percent. The best-fit model is incapable of matching the observed numbers exactly, but it reproduces the trend and approximates the values.

Figure 6.11 shows the results for cohort retention. The "observed" curve is the cohort retention curve (i.e., the cohort retention rate as a function of time since initiation) determined by averaging the cohort retention curves calculated from three different years of the NHSDA (1985, 1988, and 1990). This is compared to the cohort retention curve generated by the fitted Markovian model. As required, the modeled ten-year cohort retention rate exactly matches the observed ten-year rate. Because the modeled and observed curves are characterized by approximately the same shape, we concluded that fitting to only the ten-year retention rate (and not to other N-year retention rates, where $N < 10$) was adequate.

In summary (see Table 6.1), the fitted Markovian model tracked the historical data fairly well. Prevalence over the ten survey years from 1972 to 1991 was tracked satisfactorily, as can be seen by the table's first entry, which shows that the average modeled prevalence roughly matches the average observed prevalence. The percentage of all users that are heavy users was required to follow the trend from 1985 to 1990, and to approximately match the values for the three survey years. The second line of the table shows that the heavy-percentage fit was good. The modeled ten-year cohort retention rate was required to match the observed rate. It did, as reflected by the third line of the table.

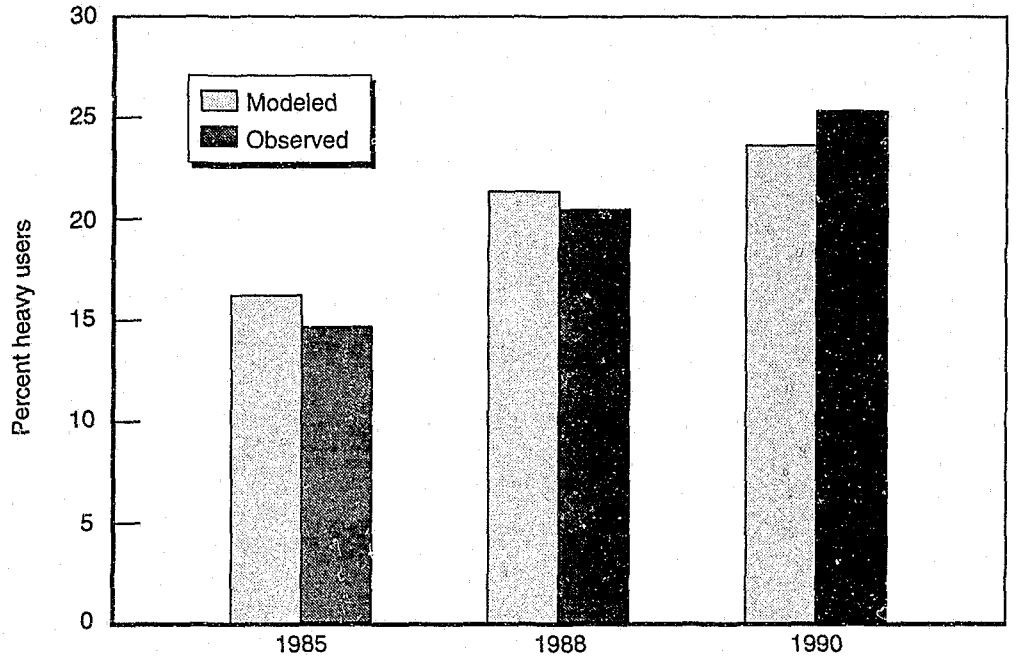

Figure 6.10—Percentage of Users That Are Heavy Users: Modeled vs. Observed

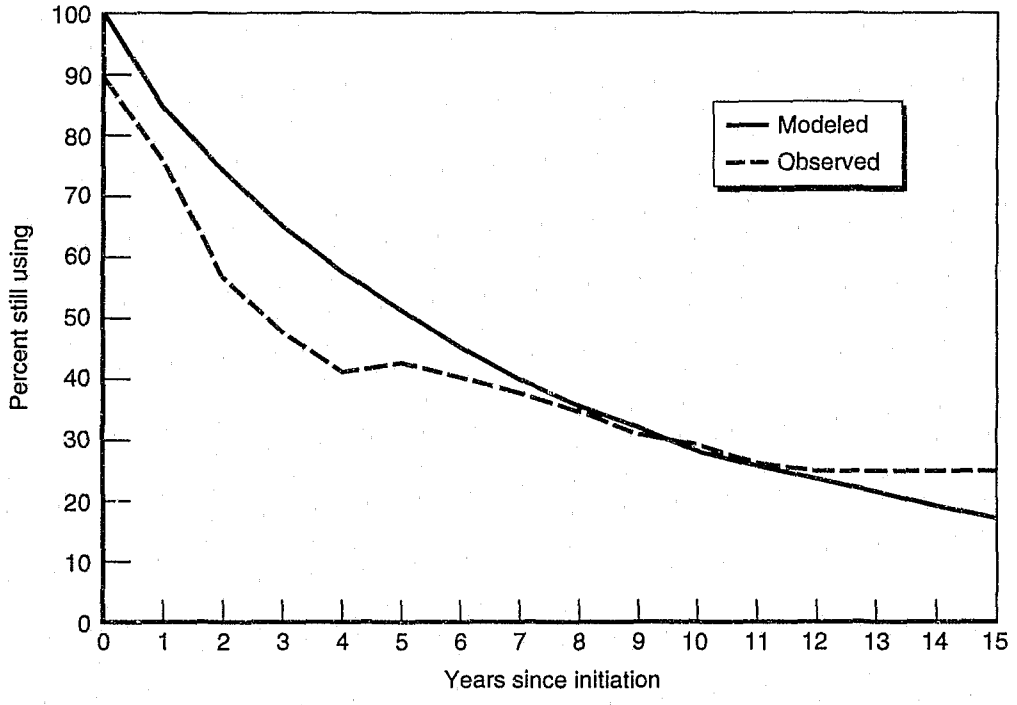

Figure 6.11—Cohort Retention: Modeled vs. Observed

Table 6.1

Summary Measures of Model Performance

Measure	Observed	Modeled
Prevalence (averaged over NHSDA survey years from 1972 to 1991)	7.2 million	6.7 million
Percent heavy users (averaged over 1985, 1988, and 1990)	20.1%	20.4%
Ten-year cohort retention	29.1%	29.0%

One valuable application of the model is to use it to explore various aspects of the cocaine epidemic history that are not obvious from examining the raw data. The graph in Figure 7.1 depicts the modeled percentage of all users that are heavy users over time. The dip in the percent-heavy-user curve just before 1980 corresponds to the rapidly increasing incidence that occurred around that time (a consequence of the fact that all new users are light users). Since that time, the percentage of all users that are heavy users has increased dramatically.

As was discussed in Chapter Five, incidence to cocaine use peaked around 1980 and has subsequently decreased (until, perhaps, very recently). Figure 7.1 suggests that *on the macro scale,* there is a delay of about ten years between incidence and heavy

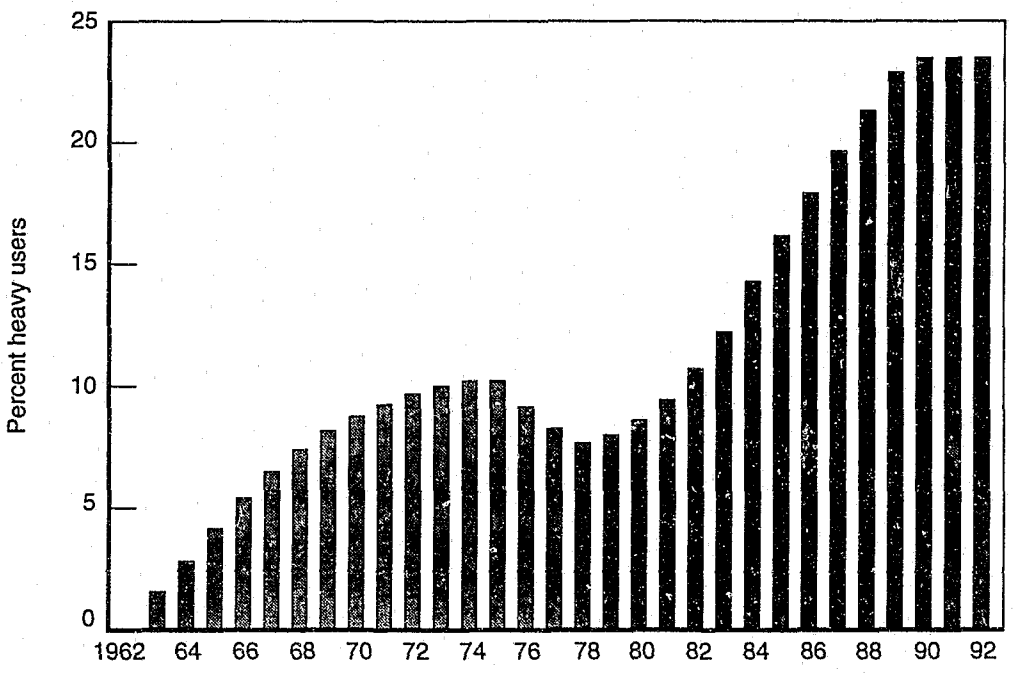

Figure 7.1—Modeled Percentage of Users That Are Heavy Users: Variation over Time

usage.[1] Consequently, the effect on heavy usage of government programs that reduce incidence (such as prevention programs) will only be realized many years later. And part of the effectiveness of local law enforcement and other programs that influence drug use in multiple ways (affecting incidence, flow rates, and the consumption rates of current users) will also be delayed. Thus, it is critical that such delays be considered when the benefits of these types of programs are measured. (Heavy-user treatment programs will affect the number of heavy-users more directly and immediately, but the effects of treatment programs are also delayed in the sense that most users need to undergo treatment regimes several times for the treatment to be "effective.")

The significant variation in the heavy-user percentage implies that overall prevalence is an incomplete and insufficient measure of the status of the cocaine epidemic. When a larger fraction of the overall prevalence is associated with heavy users, a different cocaine control strategy might be desirable. For example, prevention programs (which are hypothetically most effective in the early part of the epidemic when users are few, most of the users are light users, and potential users are many) could be scaled back while treatment programs are expanded to respond to the greater proportion of heavy users that emerge in the latter part of an epidemic.

The graph of modeled prevalence over time in Figure 7.2 reveals the underlying contributions to the prevalence estimates by light users and heavy users. Both the overall prevalence and the light-user prevalence exhibit a peak in the early part of the last decade and a more recent leveling off. In fact, this *overall* prevalence curve has characterized the course of the cocaine epidemic in the eyes of some policymakers. But while both overall and light-user prevalence have recently declined and leveled off, the number of heavy users has continued to increase. This strongly suggests that the "cocaine problem" is not disappearing (as some responsible for drug control in the government are eager to announce). To the contrary, the "problem" may be getting worse. (Of course, this depends on how "problem" is defined.)

Combining estimates of how much cocaine light and heavy users consume with the modeled prevalence information (Figure 7.2) gives a picture, displayed in Figure 7.3, of how cocaine consumption has varied over time. Here we use an estimate of 291 metric tons of cocaine consumed in 1992 (Rydell and Everingham, 1994). This total consumption estimate, the estimated (modeled) number of light and heavy users in 1992 (i.e., 5.5 million light and 1.7 million heavy users at the start of 1992), and the ratio of heavy- to light-user annual consumption (calculated in Chapter Three as about 8:1) can be combined algebraically to determine that, on the average, light users consume about 16.4 grams per year and heavy users consume about 118.9 grams per year.[2]

---

[1]Note that this is not necessarily true for an individual user; this false conclusion is an example of the fallacy of division.

[2]These averages include the incarcerated population, which consumes only negligible amounts of cocaine. The average light and heavy consumption rates for the cocaine-using populations are 17.2 and 140.0 grams per year, respectively.

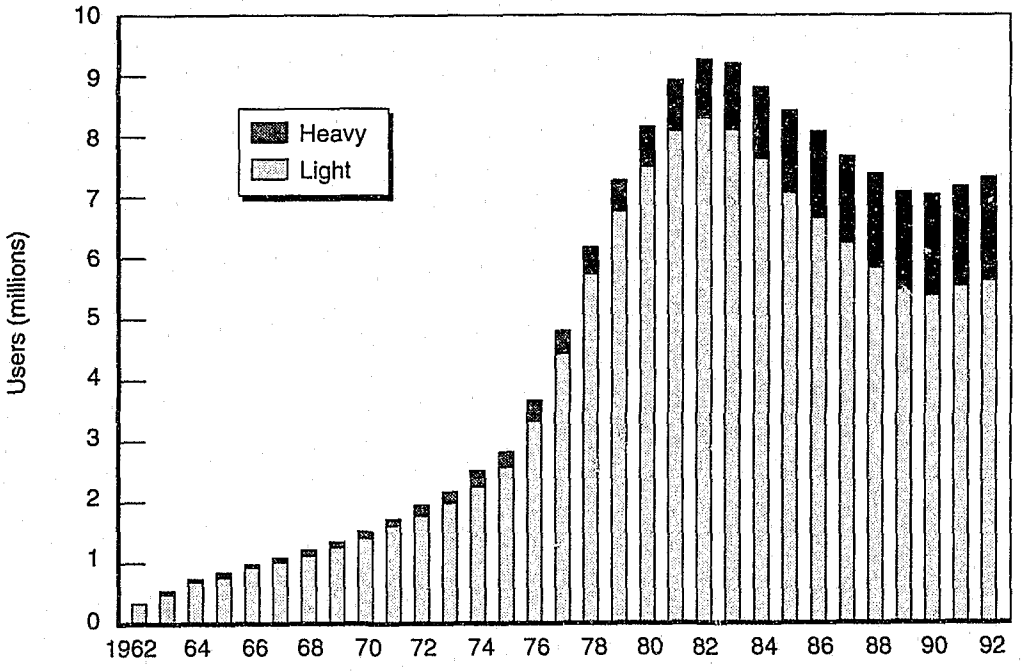

Figure 7.2—Modeled Prevalence: Heavy vs. Light Users

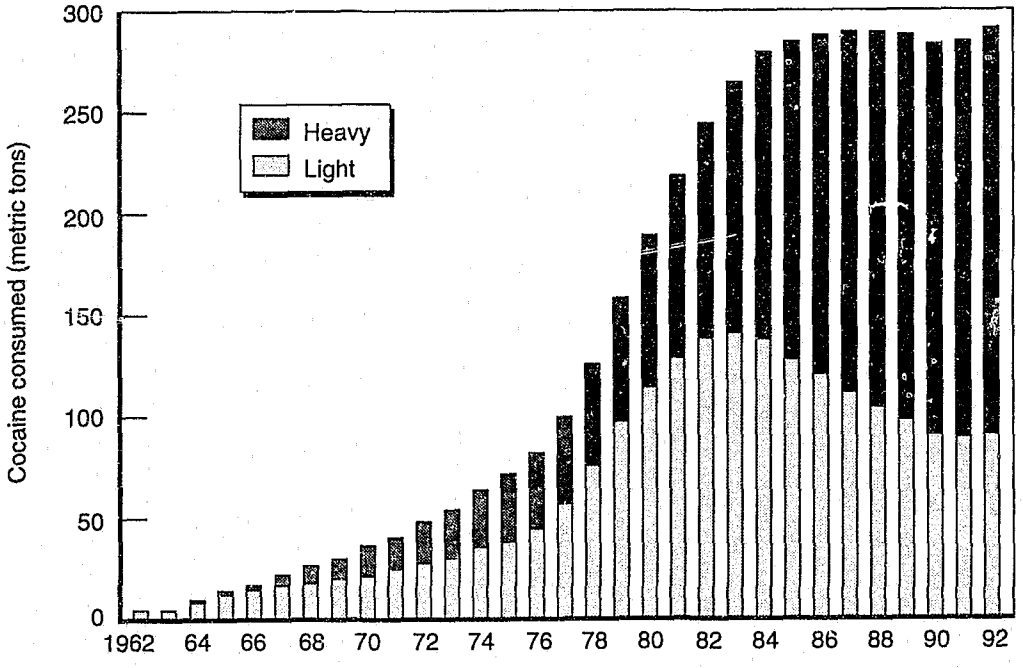

Figure 7.3—Modeled Consumption: Heavy vs. Light Users

In contrast to the overall decline in prevalence over the past decade, consumption has merely leveled off, as is evident from Figure 7.3. If incidence continues to decline, consumption will also decline, but this decline will not be noticeable for years. Even while overall prevalence is declining, large amounts of cocaine will still be consumed in the United States by the remaining users because more and more of them will be heavy users. To counter this trend, cocaine control programs that focus upon reducing consumption by heavy users are required.

The bottom line is that, not only in terms of the prevalence of heavy users, but also in terms of total cocaine consumption, the "war against cocaine" has by no means been "won." This conclusion supports those who argue that the cocaine problem is worsening. The effectiveness and costs of various cocaine control programs must be compared to determine what control strategy is optimal at this point in the epidemic.[3]

---

[3]This is the topic of the companion document to this one: Rydell and Everingham (1994).

Chapter Eight

FINDINGS: PROJECTING THE FUTURE

Because cocaine incidence is an input to the model and the course of the cocaine epidemic depends so strongly upon incidence, our model by itself is not predictive of the future course of the epidemic. However, given a script for future incidence, the model can answer certain questions about the future. For example, it can show how long it would take for the epidemic to (nearly) disappear if there were no future incidence. More generally, the model can project the course of the epidemic given any hypothetical incidence scenario.

Obviously, whether such a projection actually predicts the future course of the epidemic strictly depends on whether the corresponding incidence scenario proves to be true. But the hypothetical incidence scenarios, and the resulting prevalence and consumption projections, are much more than futile guesses destined to be wrong because future incidence cannot be predicted with any certainty. On the contrary, the value of such projections lies in the fact that they bound the analysis in a useful way.

In this chapter, the 15-year course of the epidemic is projected for a number of different incidence scenarios. For each scenario, incidence, light and heavy prevalence, and light and heavy consumption (assuming constant consumption rates) are plotted separately. Figure 8.1 shows the three graphs for the first incidence scenario, the worst case considered, for which it is assumed that annual incidence remains at the level estimated for 1991: 0.988 million new users per year.[1] From the prevalence and consumption graphs, it is evident that constant incidence, even at a magnitude as low as it has been in recent years, implies both an increase in prevalence of about 1 million users over the course of 15 years, and a substantial increase in the amount of cocaine consumed in the United States. Thus, in the absence of cocaine control programs that significantly alter the flow rates, incidence must decrease if cocaine use is to be counteracted.

Figure 8.2 shows the results for scenario 2, in which we assume incidence is halved in the next 15 years. (This is roughly equivalent to the incidence trend between 1984 and 1989—see Figure 5.1.) Halving the incidence reduces current overall prevalence by only about 1 million users (second graph) and does not reduce consumption at all (third graph).

[1] Of course, this is not the worst possible case, since incidence could once again follow an increasing trend.

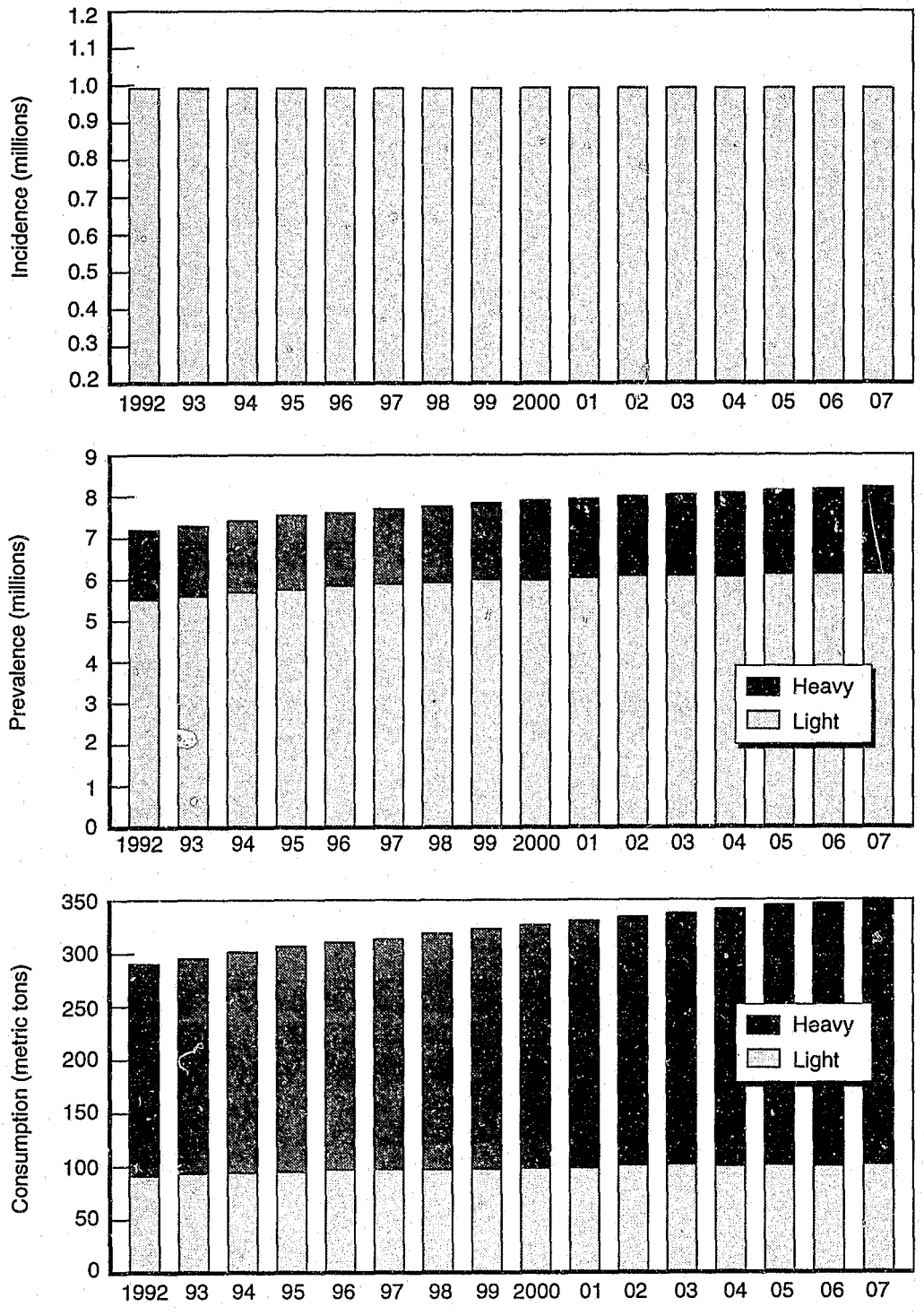

Figure 8.1—Hypothetical Scenario 1: Constant Incidence

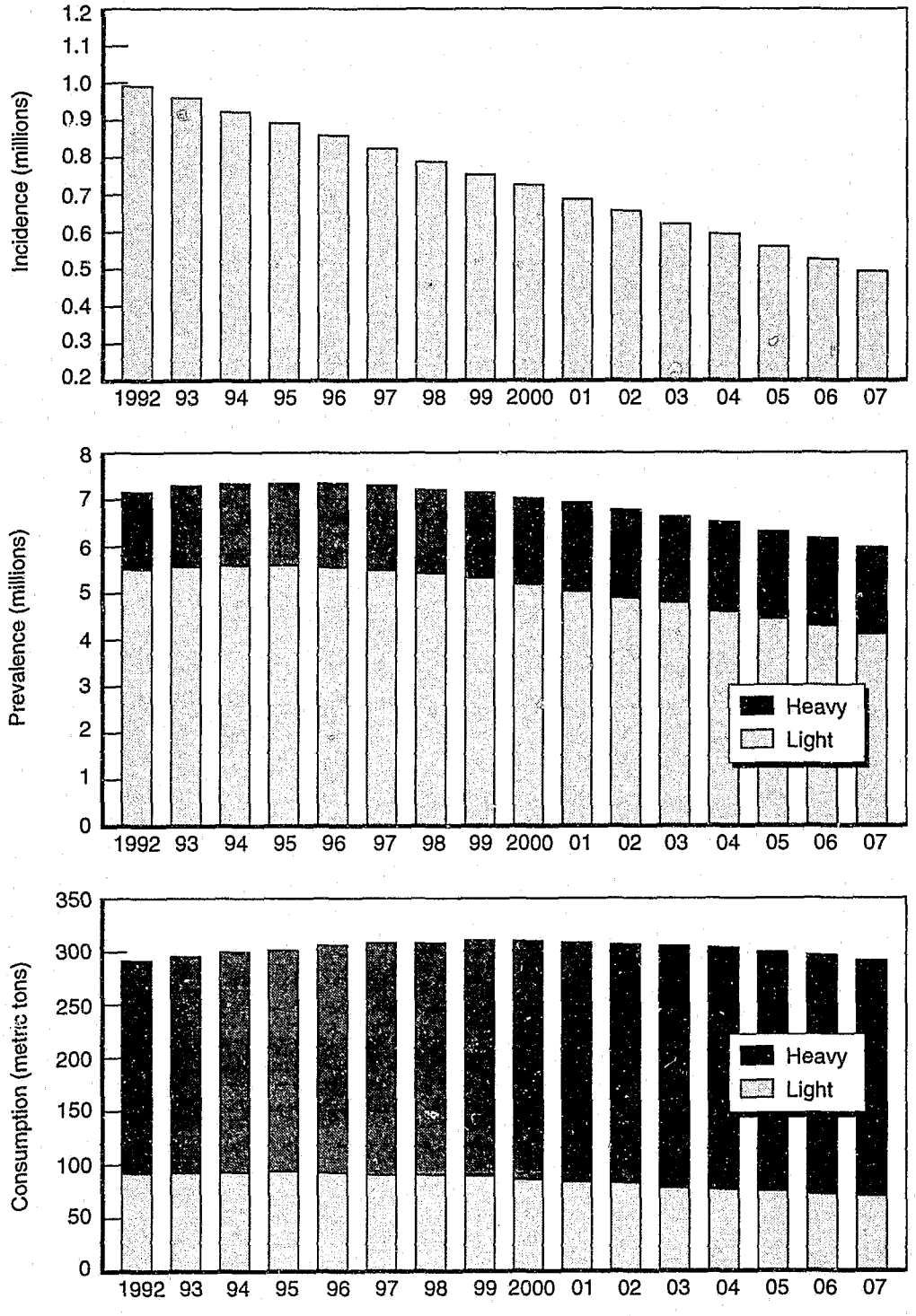

Figure 8.2—Hypothetical Scenario 2: Incidence Halved in 15 Years

A third, more optimistic scenario is plotted in Figure 8.3. This one involves an incidence decline extrapolated from the retrospective estimate of incidence (instead of the average estimate—see Chapter Five), which is near zero in 15 years. The corresponding prevalence is less than it was in the previous scenario by about 1 million users over 15 years (see second graph in Figure 8.3); however, in spite of the optimistic incidence projection, consumption decreases only marginally over the 15-year period (see third graph). This is a direct consequence of the persistence of heavy users and suggests that reducing incidence, while necessary, is by no means sufficient.

How would a sudden but temporary surge in incidence, perhaps as the result of a short-term cut in prevention funding, affect the epidemic over several years? In this fourth scenario, shown in Figure 8.4, incidence is halved over the course of 15 years (as in the second scenario), except for one year in which it is drastically increased. By comparing the prevalence and consumption graphs here with those in Figure 8.2, we see that it takes just about 15 years to recover from the temporary surge in incidence.

Having observed that a steady decline in incidence only marginally affects the course of the epidemic, one wonders what is the maximal decrease that incidence, or rather the lack thereof, could cause in prevalence and consumption. Assuming (optimistically and probably quite unrealistically) that incidence is reduced to zero and does not resurge, the maximum effect that reduced incidence can have on the future course of the cocaine epidemic can be estimated. Figure 8.5 shows the results for this fifth scenario. As the second graph shows, prevalence is reduced to about 2 million cocaine users in 15 years. But, since most of those users are heavy users, the decrease in consumption (third graph) is not nearly as dramatic: in 15 years, consumption is only halved. Thus, even in the absence of incidence, it would take about 30 years for the current epidemic to (nearly) disappear, unless programs that increase the flow rates out of cocaine use are expanded.

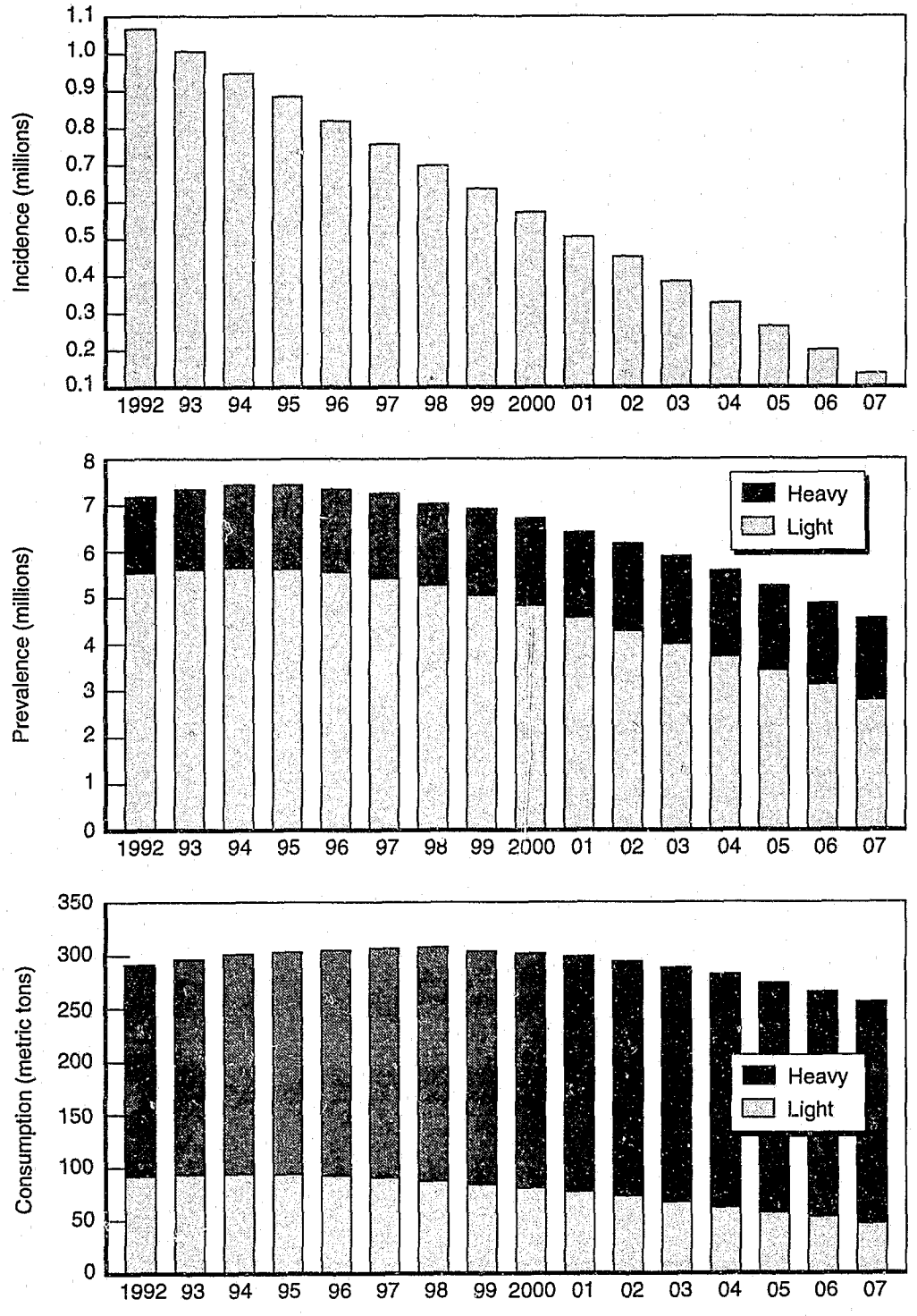

Figure 8.3—Hypothetical Scenario 3: Incidence Extrapolated to Near-Zero in 15 Years

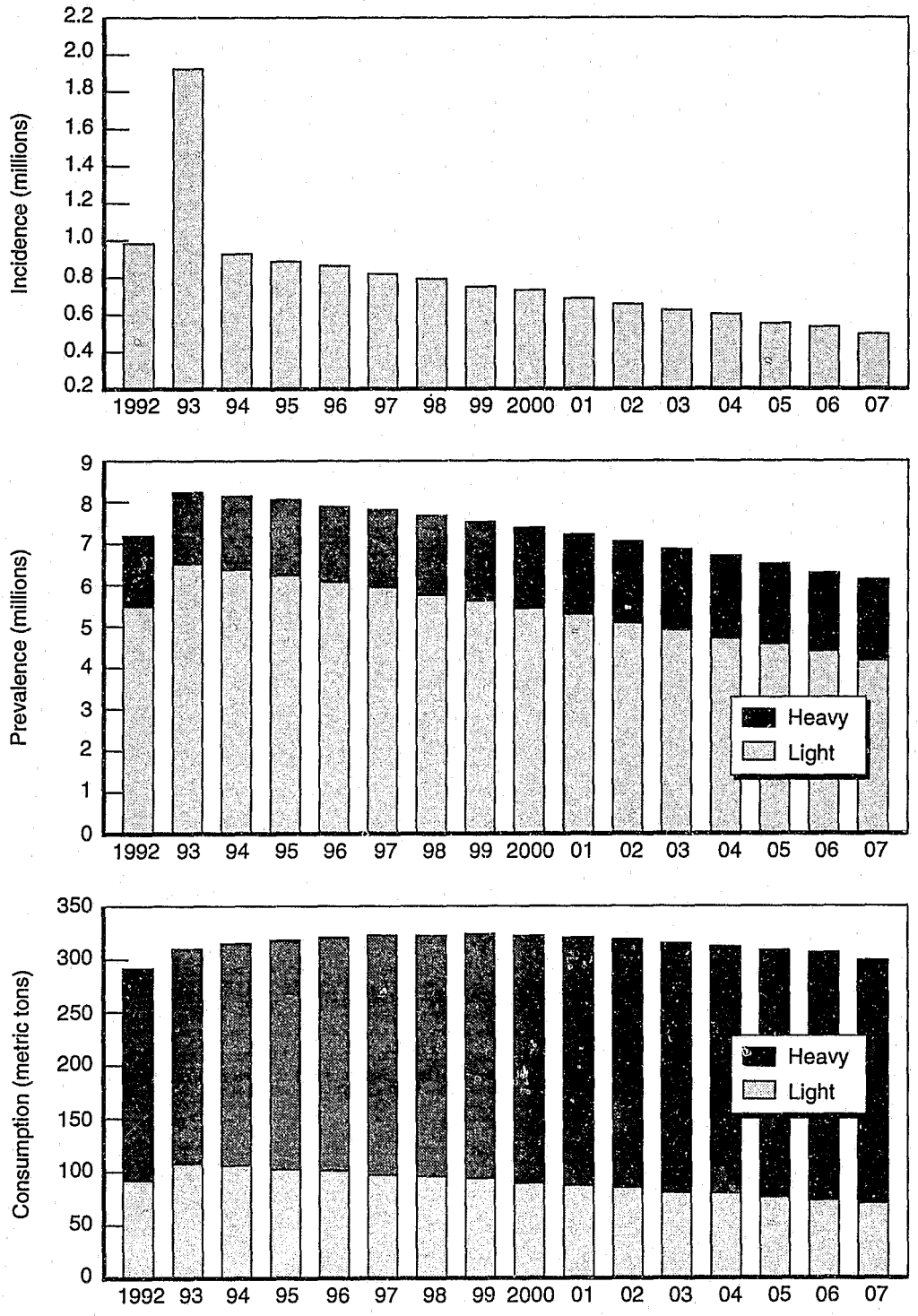

Figure 8.4—Hypothetical Scenario 4: Temporary Surge in Incidence

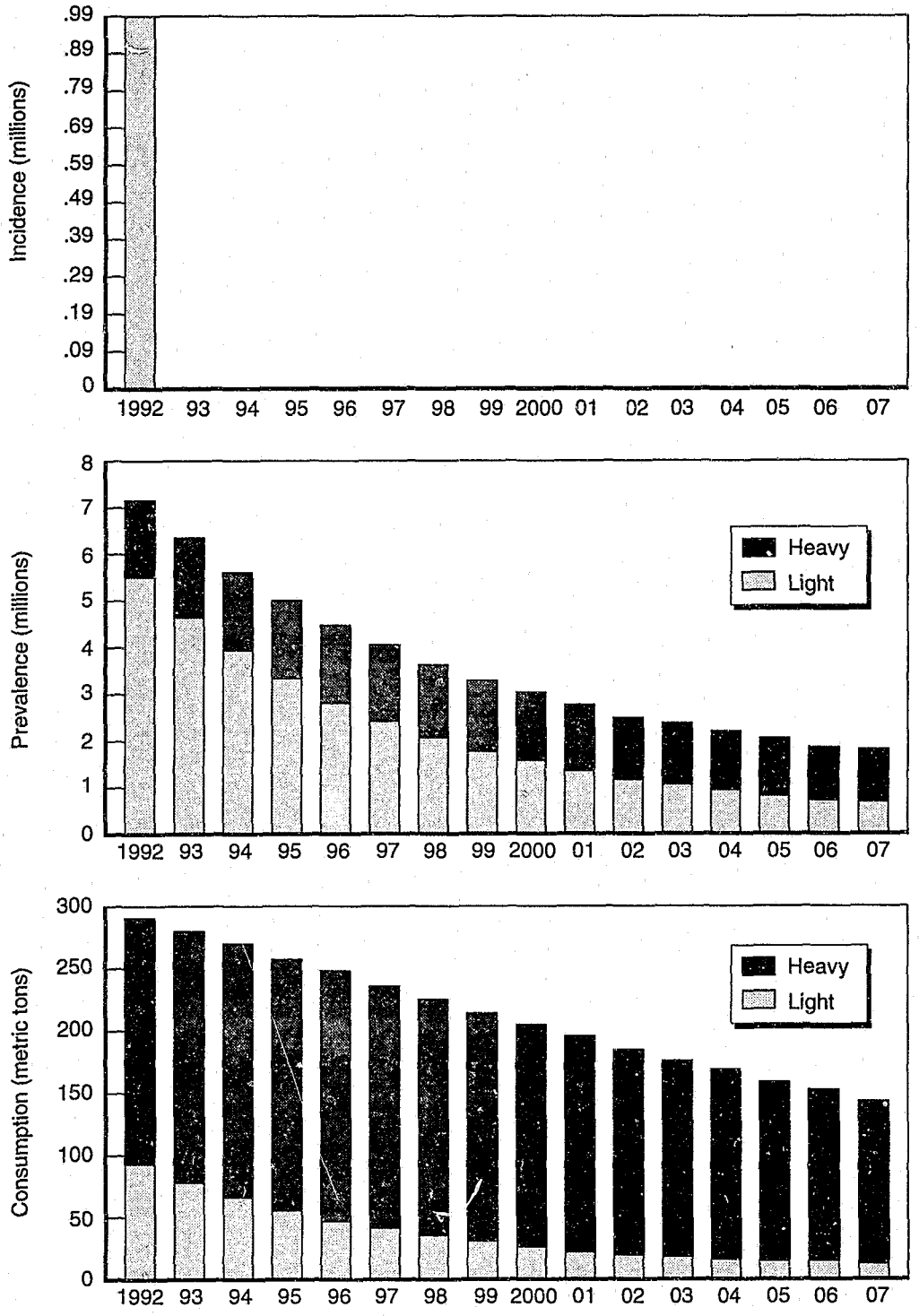

Figure 8.5—Hypothetical Scenario 5: Zero Incidence

REFERENCES

Bureau of Justice Statistics (1991), *Source Book of Criminal Justice Statistics, 1990*, Office of Justice Programs, U.S. Department of Justice.

Burt, Martha R., and Barbara E. Cohen (1989), *America's Homeless: Numbers, Characteristics, and Programs That Serve Them*, Urban Institute, Report 89-3, July.

Dombey-Moore, Bonnie, Susan Resetar, and Michael Childress (forthcoming), *System Description of the Cocaine Trade*, MR-236-A/AF/DPRC, Santa Monica, CA: RAND.

Ebener, Patricia A., Eva Feldman, and Nora Fitzgerald (1993), *Federal Databases for Use in Drug Policy Research: A Catalogue for Data Users*, N-3562-DPRC, Santa Monica, CA: RAND.

Falck, Russel, Harvey A. Siegal, Mary Ann Forney, Jichuan Wang, and Robert G. Carlson (1992), "The Validity of Injection: Drug Users Self-Reported Use of Opiates and Cocaine," *The Journal of Drug Issues*, Vol. 22, No. 4, Fall, pp. 823–832.

Fischer, Pamela J. (1987), "Substance Abuse in Contemporary Homeless Populations," Alcohol, Drug Abuse, and Mental Health Administration Conference on Research Methodologies Concerning Homeless Persons with Serious Mental Illness and/or Substance Abuse Disorders, July.

Gerstein, Dean R., and Henrick J. Harwood, eds. (1990), *Treating Drug Problems, Volume 1: A Study of the Evolution, Effectiveness, and Financing of Public and Private Drug Treatment Systems*, National Academy Press.

Gfroerer, Joseph, and Marc Brodsky (1992), "The Incidence of Illicit Drug Use in the United States, 1962–1989," *British Journal of Addiction*, No. 87, pp. 1345–1351.

Harrison, Lana D. (1991), "Methodological Design of the National Household Survey on Drug Abuse in the USA," Second Conference on Drug Use and Drug Policy, Lyon, France, September 26–28.

Homer, Jack B. (1990), *A System Dynamics Simulation Model of Cocaine Prevalence*, UCLA Drug Abuse Research Group, February.

Homer, Jack B. (1993), "A System Dynamics Model for Cocaine Prevalence Estimation and Trend Projection," *The Journal of Drug Issues*, Vol. 23, No. 2, Spring, pp. 251–279.

Hser, Yih-Ing (1993), "Prevalence Estimation: Summary of Common Problems and Practical Solutions," *The Journal of Drug Issues*, Vol. 23, No. 2, Spring, pp. 335–344.

Levin, Gilbert, Edward B. Roberts, and Gary B. Hirsch (1975), *The Persistent Poppy: A Computer-Aided Search for Heroin Policy*, Ballinger Publishing Company.

Mieczkowski, Thomas (1990), "The Accuracy of Self-Reported Drug Use: An Evaluation and Analysis of New Data," *Drugs, Crime and the Criminal Justice System*, Anderson Publishing.

Musto, David F. (1973), *The American Disease*, New Haven, CT: Yale University Press.

National Institute on Drug Abuse (1991), "Overview of the 1991 National Household Survey on Drug Abuse," press release, December.

Office of National Drug Control Policy (1992), *National Drug Control Strategy*, Washington, DC: The White House, January.

Rhodes, William (1993), "Synthetic Estimation Applied to the Prevalence of Drug Use," *The Journal of Drug Issues*, Vol. 23, No. 2, Spring, pp. 297–321.

Rydell, C. Peter, and Susan S. Everingham (1994), *Controlling Cocaine: Supply Versus Demand Programs*, MR-331-ONDCP/A/DPRC, Santa Monica, CA: RAND.

U.S. Department of Health and Human Services (1991), "1991 National Household Survey on Drug Abuse," press release, December 19.

Wickens, Thomas D. (1993), "Quantitative Methods for Estimating the Size of a Drug-Using Population," *The Journal of Drug Issues*, Vol. 23, No. 2, Spring, pp. 185–216.