Cocaine Use in America: The Evolution of a Dynamic Model

Jack B. Homer
University of Southern California
Los Angeles, California 90089-0021

in association with the research staff of the
Drug Abuse Research Group
University of California, Los Angeles

ABSTRACT

This paper describes the development of a System Dynamics model of cocaine use in the United States of America. The model's evolution is presented chronologically as a story in which theory and data have interacted and changed over time. This story may be particularly instructive for those System Dynamics modelers working, under conditions of some change and uncertainty, on extended studies of social behavior. An approach which combines skepticism, flexibility, and attention to detail throughout such studies is advocated. When a variety of alternative theories and hypotheses is available, as in many social science applications, it is important to gather a wide spectrum of relevant evidence in order to reduce the risk of model misspecification and improve the study's effectiveness.

INTRODUCTION

In 1987, I was invited to join a team of drug abuse researchers at UCLA in writing a proposal to the National Institute of Justice (NIJ) for a two-year study which "improves upon existing models and develops new models to be used for estimation of drug use prevalence". System Dynamics was presented as one of three modeling techniques (the other two, Multiple Recapture Census and Synthetic Estimation, being based in statistical theory) which had been previously used for heroin prevalence estimation and which, in our study, would be applied to estimating and projecting into the near future the prevalence of cocaine use in the United States.

Drug use prevalence estimates are relied upon by a host of governmental agencies for enforcement planning, criminal justice system needs, and resource allocation for prevention and treatment efforts. The recently established Office of National Drug Control Policy has also relied upon such estimates in carving out national priorities in the "war on drugs", and these numbers may affect the fate of legislation aimed at reducing the use and supply of illicit drugs.

One may ask why models are needed for drug prevalence estimation, when population surveys performed by the National Institute on Drug Abuse (NIDA), including the National Household Survey on Drug Abuse (NHS) and the High School Senior Survey (HSSS), would appear to serve this function. Indeed, the results of these surveys are typically presented with great fanfare and often treated by policymakers as officially "blessed". However, recognition has been growing that these surveys do not tell the entire story, and in particular that they may tend to underrepresent or exclude certain "hidden" portions of the population among whom illicit drug use is particularly prevalent. (For example, the NHS samples only people living in households, excluding such population groups as the incarcerated and the homeless.) Also, trends apparent in population survey results may not always appear to be consistent with trends in other indicators of illicit drug use, such as data from hospitals on drug-related emergency room episodes and deaths (assembled by NIDA's Drug Abuse Warning Network, or DAWN). In addition, population surveys (the
NHS, for example) may be criticized for not providing results on a timely or regular basis and for not providing a clear basis for making projections.

System Dynamics was presented in our 1987 proposal as a technique capable of (1) encompassing and relating a wide variety of relevant indicator data, (2) providing continuous-time estimates of specified user categories of interest (from 1976 to the present), and (3) providing near-future trend projections (through 1992). The study was funded at the end of 1987, and a final report has recently been submitted (UCLA 1990). Our team has now been awarded a third year of funding to extend the work, which includes refocusing the System Dynamics model to emphasize cocaine-related crime and handle a variety of policy issues.

This paper describes the development of the model through the present time, broken down into the following five periods of activity:

1. Literature Review and Initial Dynamic Hypothesis [10/87 - 12/87];
2. Initial Data Collection and Modeling [1/88 - 6/88];
4. Model Refinement and Documentation [12/88 - 6/89];
5. Further Model Refinement Toward Publication [7/89 - 4/90].

LITERATURE REVIEW AND INITIAL DYNAMIC HYPOTHESIS [10/87 - 12/87]

A review of the literature bearing on the modeling of drug use prevalence had started during the writing of our proposal and was continued during the study's early months. Although we continued to stay abreast of written reports and news of progress by other researchers throughout the study, these first few months of investigation provided a rich background for beginning the task of conceptualization, which culminated in an initial dynamic hypothesis.

Bearing most directly on this initial conceptualization effort were two previous system dynamics modeling studies of illicit drug use. The first is the "Persistent Poppy" model (Levin, Roberts & Hirsch 1975), which examines heroin use in New York City from the standpoint of policy rather than prevalence estimation. This relatively large model contains a wealth of endogenous factors of concern to urban decision makers, including local heroin availability and price, the extent of "drug culture", law enforcement activity, and educational and treatment programs; even the community's socioeconomic status was made endogenous in some Persistent Poppy simulations. Also, the model includes a detailed structure of population stocks and flows, following individuals from general drug use to heroin use, and from street-level addiction to treatment, prison, and community rehabilitation. Unfortunately, the model was developed at a time when the numerical data needed for calibration and validation were lacking, leaving the relative significance of its many components open to question.

Much simpler than the Persistent Poppy is a family of models by Gardiner and Shreckengost addressing the issue of drug supply and demand on a national level. These models were originally designed to look at heroin (Gardiner & Shreckengost 1985, 1987) and later applied to cocaine (Shreckengost 1985). The central variable is "relative abundance", a comparison of supply with actual or potential demand which is defined one way in the heroin models and another way in the cocaine model. The models have been used primarily to make inferences about the flow of imports, which is modeled as changing exogenously over time. With the imports variable adjusted appropriately, both the heroin and cocaine models have been able to reproduce certain historical indicator time series, including the past-year user population (as reported by the NHS) and retail price and purity.

From the standpoint of prevalence estimation, both of the previous system dynamics studies were seen as having something to offer. The supply-and-demand model in particular seemed appealing
for its relative parsimony and its proven ability to reproduce national indicator data series; and we noted that price and availability of supply are seen by many researchers as important variables affecting illicit drug use (Gold, Washon & Dackis 1985; Grabowski & Dworkin 1985; Siegel 1985). In addition, Shreckengost had already devoted some effort to estimating the number of heroin users with a slightly modified version of the model in which users (an auxiliary variable) are disaggregated into separate categories of light, medium, and heavy use (Shreckengost 1984a). But he found that his estimates were rather sensitive to the definitions of these user categories, and that historical increases in the number of heroin addicts were not fully explained by relative abundance alone (Shreckengost 1984b). We speculated that this shortcoming of the supply-and-demand model might be overcome by explicitly modeling the process of escalation from casual use to compulsive use, thereby giving the model some endogenous "momentum" in its portrayal of users. This hypothesis would imply incorporating a structure of population stocks and flows in the manner of (though not so complex as) the Persistent Poppy model.

Moreover, we realized that a model for estimating and projecting the prevalence of use could not very well be driven by a variable, imports, for which no truly reliable and continuous estimates, let alone projections, are available. Therefore, our initial dynamic hypothesis included linkages from relative abundance not only to demand, as in the Gardiner and Shreckengost models, but also to supply. In particular, we considered portraying an entire chain of supply – starting with coca cultivation in South America, traveling through a system of production and distribution, and ending with retail sales – which is sensitive to relative abundance in the U.S. and is capable of influencing the drug's prevalence over a period of years.

INITIAL DATA COLLECTION AND MODELING [1/88 - 6/88]

During the first half of 1988, we pressed forward in both data collection and model building. Data were collected from NIDA and the National Narcotics Intelligence Consumers Committee (NNICC) and included the following:

- HSSS prevalence estimates annually from 1976 to 1987;
- DAWN cocaine-related morbidity and mortality data annually from 1977 to 1985;
- NNICC estimates of imports annually from 1981 to 1985;
- NNICC estimates of federal seizures annually from 1977 to 1986;
- NNICC estimates of price (retail and wholesale) and purity annually from 1980 to 1986.

The NHS and the HSSS include questions regarding drug use during the past month, during the past year, and during the individual's lifetime. Figures 1 and 2 present, respectively, the past month and past year estimates, expressed as fractions of the total population represented by each survey (updated through 1989); the NHS population includes all household members aged 12 and older, while the HSSS population includes only high school seniors (not including dropouts or absentees). In these two figures, one can discern trends in use starting with rapid growth through 1979, followed by gradual decline through 1983, followed by some growth through 1985, followed by decline in the late 1980's.

DAWN data on emergency room mentions are presented in Figures 3 and 4 (updated through 1988). These data, even if looked at only through 1985, appear to show trends somewhat different from those seen in the population survey data. One sees a pattern of gradual growth in morbidity through 1983, followed by the beginnings of accelerated growth in 1984 and 1985. This accelerated growth in morbidity coincides with growth in the fraction of emergency room mentions in which smoking was the reported route of administration, rather than sniffing or injecting the drug (or other miscellaneous routes). This suggested to us the growing significance of "crack", an easily transported form of smokable cocaine which was first reported in Southern California and
Texas in 1981, spread to New York City in 1984, and was found in urban areas all over the
country by 1986 (NNICC 1987; Johnson et al. 1987).

NNICC's estimates of imports and federal seizures of cocaine (data not reproduced in this paper)
are based primarily on information from the Drug Enforcement Administration (DEA), which has
field offices throughout the U.S. and is involved in international monitoring and enforcement
activities. The imports estimates are quite rough and derived largely from estimates of global coca
cultivation and production capacity. The data on seizures are considerably more accurate and come
from the DEA's domestic field offices as well as from the Coast Guard and Customs Service.
Both the imports and seizures estimates seemed to portray a trend of continuous growth very
similar to that seen in the DAWN data.

The NNICC estimates (low-end and high-end) of retail price and purity for 1980-1986 are bar-
graphed in Figures 5 and 6, respectively. We learned that these estimates might be unreliable,
largely because they represent the subjective consensus of a committee rather than consistent
information from a single data base. Shreckengost had, in his cocaine study (1985) independently
investigated the DEA's STRIDE (System To Retrieve Drug Evidence) data base on domestic
seizures, and, filling in certain gaps in this data base with subjective estimates, found a pattern of
continuously declining price and continuously increasing purity from 1975 to 1984. (We later
obtained the STRIDE data tapes so that we could perform our own investigation, as discussed
below, which yielded the price and purity figures shown as line-graphs in Figures 5 and 6.)

As numerical data and other sorts of evidence were accumulating, our initial dynamic hypothesis
was refined and converted into a testable model. Figure 7 presents an overview of the model as of
June 1988. Throughout the first half of 1988, we continued to believe that the interplay of supply
and demand, combined with the momentum of user population flows, would be sufficient to
explain the available evidence. We hypothesized that, while the NHS and the HSSS may provide a
reasonable picture of total use, the DAWN and NNICC indicators were telling us much more about
the growing (yet "hidden") population of compulsive users in particular. It is generally accepted
that compulsive use is much more likely than casual use to lead to medical difficulties (Gawin &
Kleber 1985; Gold, Washton & Dackis 1985; Siegel 1985). Also, compulsive users consume a
disproportionately large amount of the drug. For example, Shreckengost (1985) has estimated
that, in 1982, daily-to-weekly users of cocaine were responsible for 86% of total consumption
though they made up only 19% of total users that year. Thus, it is quite conceivable that medical
emergencies and drug supply might grow while the total number of users does not.

One area of model refinement during early 1988 was in the area of population stocks and flows.
Where initially we had categorized users only by intensity of use (casual or compulsive), we now
added the dimension of recency of use (past month, past year but not past month, prior to past
year). This refinement was made in part to give the model more points of comparison with the
NHS and the HSSS. Also, it allowed us to represent the common phenomenon of relapse, both
within a year after discontinuation and afterwards.

Other refinements were in the area of supply. First, we decided not to portray a level of domestic
inventory, as Gardiner and Shreckengost had done in their models, based on DEA estimates that
such an inventory may comprise a few weeks' worth of consumption at most; instead, we assumed
that all unseized imports are rather quickly consumed. Second, we decided not to model foreign
production and distribution, based on information that foreign cocaine supplies and production
capacity appear historically to have been so plentiful that smugglers to the U.S. have been able
obtain as much as they would like at the same low price. Although we recognized that such sim-
plications of the supply process might compromise the model's ability to evaluate the impacts of
policies such as foreign crop eradication, we decided to err on the side of parsimony for the time
being.
In June 1988, a non-mathematical report on the existing model and its underlying assumptions was written and distributed for external review. The model had not yet been calibrated in an effort to reproduce historical data, so we did not know its adequacy in this regard. In the report, we acknowledged that the cocaine user population might be affected by factors other than price, and we singled out the social "aura" of cocaine and the availability of crack as two factors that might ultimately have to be modeled independently of price in order to explain the historical evidence. Others had previously pointed to the potential impact of aura on illicit drug use (Grabowski & Dworkin 1985), and the Persistent Poppy model had contained "appeal" variables affecting both general drug use and heroin use. In regard to crack availability, we were aware that others (Cole 1989, for example) had suggested that the spread of crack could be viewed primarily as reflecting the decline in cocaine's price during the 1980's, but we were uncertain on this point.

MODEL REVISION AND CONFERENCE PRESENTATION [7/88 - 11/88]

Our study team now began a period of intensive effort in anticipation of a conference to be held at the beginning of November. This "Interagency Conference on Prevalence Estimation Techniques" was being sponsored jointly by NIDA and NIJ, and would be attended by representatives from a number of different federal agencies. While NIJ was supporting our cocaine study, NIDA was funding a different group to apply a range of analytic tools – including System Dynamics – to the study of heroin. We had heard that the NIDA-supported group was limiting the System Dynamics part of their study to a reevaluation of the Gardiner and Shreckengost model, rather than attempting any original modeling of their own. Our group, in contrast, hoped to make a persuasive case for the development of new System Dynamics models for purposes of prevalence estimation.

We realized that it was now imperative to calibrate the model, to bring theory and data fully together so that the conference could see a new model which successfully reproduced and explained a broad spectrum of historical evidence. In September, we were able to obtain STRIDE tapes covering the period 1977-1987, which include quantity, price, and purity data on every undercover purchase and seizure reported to DEA field offices nationwide. After examining the full database consisting of more than 63,000 records, we restricted our "retail" analysis to include only those records where some non-zero price was recorded and where the buy was one of less than six grams. We restricted our "wholesale" analysis to include only those records where some non-zero price was recorded and where the buy was one of more than 100 grams.

The STRIDE data (in conjunction with other available data) were used to several different hypotheses contained in the June 1988 model (refer to Figure 7). These hypotheses were tested through statistical regression where possible, and otherwise through partial-model testing (Homer 1983). This testing caused us to reach the following conclusions:

1. As long as the expected profitability of smuggling is large enough (as it appears to have been throughout the 1970s and 1980s), imports are best viewed as being driven by demand, rather than as a driver of demand;

2. The wholesale to retail mark-up (see Reuter, Crawford & Cave 1988) of a pure gram of cocaine, rather than being constant as hypothesized, declined steeply during the 1980’s for reasons which are unclear, leading us to suspend attempts to model retail price endogenously;

3. The trend in retail price (after adjustment for inflation and purity) can be described as continuous but decelerating decline for the period 1977-1986, which does little to explain the complex historical patterns of cocaine use prevalence seen in Figures 1 and 2.

Having been convinced that the supply-and-demand theory of cocaine use was inadequate to the task of reproducing historical evidence, we turned next to sociological factors that might account for the drug’s social aura. For evidence of changes in cocaine’s aura over time, we examined data on attitudes, beliefs and the "social milieu" from the HSSS (see Johnston, O’Malley, and Bachman 1989) and found some interesting trends there. In particular, our attention was drawn to measures
providing evidence of changing social exposure to cocaine as well as evidence of changing perceptions of cocaine's health risks. Figure 8 presents two measures of non-exposure, whose patterns are quite similar to one another and reminiscent (if inverted) of the prevalence patterns seen in Figures 1 and 2. Figure 9 presents two measures of perceived health risk, one of them regarding experimentation and one of them regarding regular use. Of the two, we regarded the perceived health risk of regular use, which has increased more or less steadily over time, as being the more conceptually distinct variable and as having a more pervasive influence on initiation, discontinuation, and relapse of use.

Previous modeling of medical product diffusion (Homer 1987) suggested how these social variables might be modeled endogenously. A basic tenet of diffusion theory is that social acceptance increases with the number of users, and we in fact found that a simple nonlinear relationship did a good job of explaining exposure (our proxy for social acceptance) as a function of prevalence of use. Such a relationship would create a reinforcing loop in the model, connecting prevalence to exposure and exposure back to prevalence through the user flows of initiation, relapse, and discontinuation.

The increasing perception of risk, on the other hand, can be seen as a cognitive response to accumulating reports and experiences that attest to the drug's dangerous effects. Using DAWN data on cocaine-related emergencies as a proxy for such "bad news", we found that another simple nonlinear relationship did a good job of explaining perceived risk (or its inverse, perceived safety, as we formulated it) as a function of cumulative emergencies. Next, we hypothesized that the frequency of emergencies was primarily related to the number of compulsive users, with some additional contribution from casual users. This chain of relationships would create a compensating loop in the model, connecting prevalence to emergencies, emergencies to perceived safety, and perceived safety back to prevalence (again through various user flows).

At the November 1988 conference, we presented a model in which population flows are affected by the endogenous variables of exposure and safety and the exogenous variable of retail price (adjusted). This model proved capable of replicating a variety of indicator time series for the period 1976-1983, and, surprisingly, it fit history best when price was assumed to have no impact at all. The post-1983 growth in cocaine indicators was explained in this model as a function of growing crack availability, which was left essentially exogenous, but which was portrayed as having effects on several of the model's user flows as well as on morbidity and mortality. The conference seemed to agree that our modeling work was eye-opening and a move in the right direction, and that we should next focus squarely on the dynamics of crack use.

MODEL REFINEMENT AND DOCUMENTATION [12/88 - 6/89]

By April 1989, we had come to the conclusion that an appropriate representation of crack would require further disaggregation of the user population, namely by distinguishing users who prefer crack from those who prefer cocaine powder. Part of the rationale for such disaggregation is that crack appears to have opened up an entirely new market for cocaine among certain groups, particularly the black urban underclass (see Cole 1989). (Figure 4 provides some evidence for this, showing steady growth in the black population fraction of DAWN emergencies during 1984-1988.) In this revised model, individuals could either initiate directly into crack use or switch from powder to crack as their preferred form. The growth of crack use is set into motion by moving a tiny fraction of users from powder to crack in the year 1981. From that time forward, any increases in the number of crack users are assumed to lead to greater availability of (i.e., exposure to) crack, which, in turn, leads to increased attraction of crack users through initiation and switching of product preference.
The revised model did a good job of reproducing indicator data for the period 1976-1986 and was used to make projections through the year 1992, as had been promised to NIJ. It suggested that the number of users preferring powder cocaine would continue its long decline (which had started around 1980), while the number of crack users would remain large for at least several years due to crack's notorious addictiveness (see Gawin & Ellinwood 1988). It also demonstrated how it was possible for DAWN emergencies to skyrocket due to the increase in risky crack use, while the total number of users levelled off and even declined.

In June 1989, a report was written documenting the model and its assumptions in detail and presenting the results of a variety of sensitivity tests (some of them with possible policy implications) which had been performed. This report was distributed for review and later modified to become the System Dynamics portion of the final report to NIJ of our two-year study.

FURTHER MODEL REFINEMENT TOWARD PUBLICATION [7/89 - 4/90]

During the summer of 1989, reviews of the draft report and the release of new data (from the NHS, the HSS, and DAWN) led us to the realization that further model refinement would probably be required to make our work credible to an audience of researchers and decision makers in the area of illicit drug use. The existing model was unable to anticipate the rapidity of decline in reported use that had occurred during 1986-1988 (see Figures 1 and 2) as well as the rapidity of increase in cocaine-related emergencies and deaths during this same period (see Figure 3). Also, the model estimated a crack preference fraction during the late 1980's that appeared excessive based on available data. In addition, reviewers were troubled by the lack of a price effect in the existing model.

Solutions to these problems appeared to lie in a combination of parameter adjustment and structural modification. In the area of parameter adjustment, we saw the possibility that by increasing significantly the assumed morbidity and mortality risks of crack (risks for which no good independent estimates are available), the model's ability to reproduce both DAWN data and data on the fraction of crack users could be improved. The model's underestimate of the decline in reported use during 1986-1988 was unlikely to yield to such an easy fix, however, and appeared to call for structural modification.

One explanation put forward for the model's underestimated decline in reported use is that the underrepresentation biases of the NHS and the HSS have increased in recent years. This argument says that it is not so much that cocaine use has declined sharply as that it has shifted, along with the increase in crack use, toward a population that is less likely to be counted in the NHS or the HSS. A second plausible explanation put forward was that the existing model, in focusing on the aura surrounding cocaine specifically, had ignored the influence of attitudes toward illicit drug use in general. Both the NHS and the HSS track marijuana use (see Figure 10), which one might view as a guide to changes in such general attitudes, and the HSS additionally tracks the use of "any illicit drug" (see Figure 11). Figures 10 and 11 indicate that past month use of any illicit drug is highly correlated with past month marijuana use, and that a decline in illicit drug use evident throughout the 1980's has been particularly rapid since 1986.

In March 1990, experiments were performed with a somewhat simplified model specific only to the high school aged population, to determine whether a model containing the variety of factors seen in Figure 12 might be capable of reproducing HSS-reported prevalence data for the period 1976-1989. The HSS underreporting biases for past month, past year, and lifetime use were based upon previous estimates (Johnston, O'Malley, & Bachman 1989), along with our own adjustments for the crack preference fraction. This model not only proved capable of reproducing the full range of data, but it did so with parameter values which appeared reasonable in all cases. Social exposure, perceived safety, marijuana use, and retail pure price all appeared to make signifi-
cant contributions to explaining the history of reported prevalence among the high school aged population, while crack-related increases in the underreporting bias appeared to have less relative significance.

Based upon this positive result, we expect to return to the full population model to explore the implications of incorporating marijuana use and retail pure price as factors affecting actual use, as well as including a variable underreporting bias, as seen in Figure 12. We are hopeful that this proposed model will lead to publications in the illicit drug use literature, and also form the basis for further model development during our third year of NIJ funding.

CONCLUSION

Two and a half years ago, our study team speculated that a relatively straightforward modification of a previously developed System Dynamics model would prove sufficient to the task of reproducing and explaining, in a credible way, various national indicator data series concerning cocaine use in the United States. As we explored various information sources, it became increasingly clear that this speculation was not supported by the evidence and that a new and rather different model would be required instead. Our search for a more capable dynamic theory led us to explore additional sources of information and to reconsider some of our original sources. Additional model refinement was prompted when model projections were found in some instances to be inconsistent with updated information.

The story of model evolution presented in this paper may serve, in large part, to underscore conclusions reached previously by others (Randers 1973, and Richardson & Pugh 1981, for example) regarding the iterative nature of modeling. But it also says something about an approach—consisting of skepticism, flexibility, and an attention to detail—that should be maintained throughout a modeling study to maximize its effectiveness. Recognizing this as the ideal attitude of any research scientist, regardless of field, does not make it any easier to achieve.

It is of particular importance in social science modeling that one carefully consider a wide spectrum of evidence and alternative hypotheses, and continue to do so as the study progresses. When the "physics" of a dynamic system are less than obvious, or when the system is so complex that one must choose among various options for simplified representation, one should seek out all available evidence, even if this means spending valuable time analyzing voluminous data files or surveying knowledgeable experts. Only in this way can one reduce the available "degrees of freedom" for modeling and reduce the risk of model misspecification. The more that actual evidence is taken into account and the more empirically grounded the model is made, the greater becomes the probability that one's efforts will lead to a model that provides its audience with valid insights.

REFERENCES


Figure 3. Cocaine-Related Emergency Room Mentions

Figure 4. Two Key Fractions of Cocaine E.R. Mentions
Figure 5. Retail Price of Cocaine

Figure 6. Retail Purity of Cocaine
Figure 7. Overview of Initial Model (6/88)
Figure 8. Reported Non-Exposure to Cocaine

Figure 9. Perception of Cocaine as Great Health Risk
Figure 10. Past Month Marijuana Use Fractions

Figure 11. Past Month Use Fraction of "Any Illicit Drug"
Figure 12. Overview of Proposed New Model (3/90)