

Predicting Heavy Drug Use

Results of a Longitudinal Study, Youth Characteristics
Describing and Predicting Heavy Drug Use by Adults

F E B R U A R Y 2 0 0 4



EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF NATIONAL DRUG CONTROL POLICY

Predicting Heavy Drug Use

**Executive Office of the President
Office of National Drug Control Policy
Washington, DC
20503**

February 2004

Acknowledgements

This publication was sponsored by the Executive Office of the President, Office of National Drug Control Policy (ONDCP). Sam Schildhaus, Yoku Shaw-Taylor, Steven Pedlow, Michael R. Pergamit, and other staff of the National Opinion Research Center (NORC) at the University of Chicago produced this publication under a contract to ONDCP, contract number 282-98-0024, Task Order number 3. June S. Sivilli served as Project Officer for this study and resulting publication.

Disclaimer

The opinions expressed herein are the views of the authors and do not necessarily reflect the official policy of the Federal Government.

Public Domain Notice

All material appearing in this report is in the public domain and may be reproduced without permission from ONDCP or the authors. Citation of the source is appreciated. Suggested citation:

Office of National Drug Control Policy (2004). *Predicting Heavy Drug Use*
Washington, DC: Executive Office of the President (Publication No. NCJ 208382)

Electronic Access to Publication

This document can be accessed electronically through the following World Wide Web address:

<http://www.whitehousedrugpolicy.gov>

Originating Office

Executive Office of the President
Office of National Drug Control Policy
Washington, DC 20503

Publication Number NCJ 208382

Printed in January 2005

Executive Summary

Dynamic Picture of Marijuana and Cocaine Use from Age 14 to 40

- This study's main aims are to describe the movement of adolescents and young adults into and out of drug use and to predict heavy drug use. The data source is the Department of Labor's National Longitudinal Survey of Youth, which began in 1979 with a sample of 12,686 adolescents aged 14–21. After 17 rounds and 19 years, the response rate in 1998 was 87 percent.

The Transient Nature of Most Marijuana and Cocaine Use

- A majority of the sample reported no drug use in the five “index” rounds (questions on drug use were asked for five years during the 15-year period between 1984 and 1998). Overall, 42 percent of individuals reported marijuana use, 19 percent reported cocaine use, and 3 percent reported crack use (crack questions were collected only between 1992 and 1998).
- Most of the persons using any drug in a given round never used the drug again or used the drug for one additional round. Overall, 19.2 percent used marijuana, 11.7 percent used cocaine, and 2.7 percent used crack in one round only.
- Very few individuals reported drug use in all five index rounds: 3.2 percent used marijuana and 0.2 percent used cocaine in every round from 1984 to 1998. For crack, 0.2 percent used in the three rounds between 1992 and 1998.
- Only 3.9 percent used marijuana and 0.6 percent used cocaine in four rounds; crack was used in two rounds by 0.6 percent.
- Almost half of marijuana users who used in *one* index round used marijuana in the next one. Only *one* of 16 non-marijuana users in one index period *used* marijuana in the subsequent period.
- About one-quarter of cocaine users who used in one index round used cocaine in the subsequent study period. Only one of 32 non-cocaine users in one index period *used* cocaine in the next one.

Early Predictors of Heavy Cocaine Use

- All other things being equal, adolescents and young adults who smoked marijuana more than 50 times at the first contact were six times more likely to become heavy cocaine users than those who did not smoke marijuana. This finding supports the suggestion that preventing adolescents and young adults from using substantial amounts of marijuana may lead to a considerable decrease in the number of future heavy cocaine users.
- Young male users were almost twice as likely as young female drug users to become heavy drug users.

- Those who had been suspended from school were one and one-half times more likely to become heavy cocaine users than those who had not been suspended from school.
- Adolescents and young adults who reported a significant amount of illegal income at round one were more than twice as likely to become heavy cocaine users as those who did not. Also, those who sold hard drugs at round one during adolescence were two times as likely to become heavy cocaine users as those who did not. Directing adolescents and young adults from criminal occupations is not only associated with lower future criminal behavior, it is likely to be associated with substantial decreases in the number of heavy cocaine users.
- Adolescents and young adults who attended religious services at least twice a month were one-third less likely to become heavy cocaine users as those who attended no religious services. Attending infrequently is associated with a one-quarter lower likelihood of becoming a heavy cocaine user. Family and community efforts to encourage religious attendance by youth may be associated with fewer heavy cocaine users in the future.

Table of Contents

1. INTRODUCTION.....	1
Background and Purpose.....	1
Organization of this Report	2
2. METHODOLOGY.....	3
Description of National Longitudinal Survey of Youth	3
Analytic Techniques	5
<i>Description of Movement into and out of Drug Use.....</i>	<i>5</i>
<i>Prediction</i>	<i>5</i>
Variables	6
Nonresponse Analysis	9
3. DRUG USE 1984–1998	11
Comparison Between Those Included/Not Included in Analysis	11
Movement Into and Out of Drug Use	14
<i>Marijuana Use in the Past Year.....</i>	<i>14</i>
<i>Cocaine Use in the Past Year</i>	<i>17</i>
<i>Marijuana and Cocaine Use in the Past Year</i>	<i>17</i>
<i>Correlates of Number of Years Used Drugs: Gender and</i>	
<i>Racial/Ethnic Differences</i>	<i>22</i>
<i>Comparisons of Drug Use Over Time</i>	<i>23</i>
<i>Comparison Between NLSY and National Household Survey</i>	
<i>on Drug Abuse (NHSDA)</i>	<i>26</i>
<i>Strengths of Longitudinal Surveys.....</i>	<i>29</i>
Summary.....	29
4. HEAVY DRUG USE.....	31
Introduction	31
Age of First Use	31
Prediction of Heavy Cocaine Use	37
Correlates of Number of Years of Drug Use	40
Summary.....	42
5. LONGITUDINAL SCREENING.....	43
Introduction	43
Screening the Sample.....	43
Selecting the Sample	44
Weighting the Sample	45
Conclusion	45
APPENDIX A	46

List of Exhibits

	Page
EXHIBIT 2.1	NLSY79 MODE AND RESPONSE RATES BY YEAR..... 4
EXHIBIT 2.2	DRUG USE FOR INDIVIDUAL YEARS AND COMBINATIONS OF YEAR 8
EXHIBIT 2.3	FOUR TESTS OF NON-RANDOM DROPOUT AMONG HEAVY COCAINE USERS 10
EXHIBIT 3.1	COMPARISONS OF SAMPLE RESPONDENTS 1984-1998 AND SAMPLE RESPONDENTS 1984-1994..... 12
EXHIBIT 3.2	COMPARISONS OF STUDY SAMPLE AND SAMPLE EXCLUDED FROM STUDY..... 13
EXHIBIT 3.3	DISTRIBUTION TREE OF RESPONDENTS WHO REPORTED SMOKING MARIJUANA IN 1984 15
EXHIBIT 3.4	DISTRIBUTION TREE OF RESPONDENTS WHO REPORTED NO MARIJUANA USE IN 1984..... 16
EXHIBIT 3.5	DISTRIBUTION TREE OF RESPONDENTS WHO REPORTED USING COCAINE IN 1984..... 18
EXHIBIT 3.6	DISTRIBUTION TREE OF RESPONDENTS WHO REPORTED NO COCAINE USE IN 1984..... 19
EXHIBIT 3.7	DISTRIBUTION TREE OF RESPONDENTS WHO REPORTED COCAINE AND MARIJUANA USE IN 1984 20
EXHIBIT 3.8	DISTRIBUTION TREE OF RESPONDENTS WHO REPORTED NO COCAINE AND MARIJUANA USE IN 1984..... 21
EXHIBIT 3.9	NUMBER OF TIMES USED MARIJUANA BY GENDER..... 22
EXHIBIT 3.10	NUMBER OF TIMES USED COCAINE BY GENDER..... 22
EXHIBIT 3.11	NUMBER OF TIMES USED MARIJUANA BY RACE/ETHNICITY 23
EXHIBIT 3.12	NUMBER OF TIMES USED COCAINE BY RACE/ETHNICITY 23
EXHIBIT 3.13	MARIJUANA USE BETWEEN ADJACENT SURVEY PERIODS 24
EXHIBIT 3.14	COCAINE USE BETWEEN ADJACENT SURVEY PERIODS..... 24
EXHIBIT 3.15	NUMBER OF YEARS USED DRUG 24
EXHIBIT 3.16	PERCENTAGE WHO USED MARIJUANA ADDITIONAL YEARS BY NUMBER OF YEARS USED MARIJUANA 25
EXHIBIT 3.17	PERCENTAGE WHO USED COCAINE ADDITIONAL YEARS BY NUMBER OF YEARS USED COCAINE..... 25
EXHIBIT 3.18	NUMBER OF YEARS USED COCAINE BY NUMBER OF YEARS USED MARIJUANA..... 27

	Page
EXHIBIT 3.19	COMPARISON BETWEEN NHSDA AND NLSY FOR LIFETIME AND PAST YEAR MARIJUANA AND COCAINE USE 1984-1998 28
EXHIBIT 4.1	HEAVY COCAINE USE AMONG THOSE WHO HAD/HAD NOT SMOKED FIRST CIGARETTE BY AGE 32
EXHIBIT 4.2	HEAVY COCAINE USE AMONG THOSE WHO HAD/HAD NOT STARTED SMOKING DAILY BY AG 33
EXHIBIT 4.3	HEAVY COCAINE USE AMONG THOSE WHO HAD/HAD NOT STARTED DRINKING 2+ A WEEK BY AGE..... 34
EXHIBIT 4.4	HEAVY COCAINE USE AMONG THOSE WHO HAD/HAD NOT USED MARIJUANA BY AGE 35
EXHIBIT 4.5	HEAVY COCAINE USE AMONG THOSE WHO HAD/HAD NOT USED COCAINE BY AGE..... 36
EXHIBIT 4.6	HEAVY COCAINE USE AMONG THOSE WHO HAD/HAD NOT USED CRACK COCAINE BY AGE..... 37
EXHIBIT 4.7	CORRELATES OF HEAVY COCAINE USE: RESULTS OF LOGISTIC REGRESSION MODEL..... 39
EXHIBIT 4.8	CORRELATES OF NUMBER OF YEARS USING MARIJUANA, COCAINE, CRACK: RESULTS OF OLS REGRESSION MODELS 41

1. INTRODUCTION

Background and Purpose

The Office of National Drug Control Policy (ONDCP) has supported several projects to learn about heavy drug users. These projects include a major study to estimate the number of “hard core” users¹ as well as analyses of the National Household Survey on Drug Abuse (NHSDA) and Drug Use Forecasting (DUF) system.² Although these analyses have estimated the number of heavy drug users, they do not predict the correlates that distinguish heavy drug users from the remainder of the population. Such correlates would be developed with models using longitudinal data (i.e., multiple years of panel data) for characteristics (demographics, attitudes, and behavior) of youth to describe and predict heavy drug use by adults.

Given the lack of longitudinal data to describe trends and infer reasons for changes, drug policy research has relied on cross-sectional data (e.g., NHSDA, Drug Abuse Warning Network, Arrestee Drug Abuse Monitoring Program/DUF, Treatment Episode Data Set, Monitoring the Future, Drug Evaluation Network System). Some cross-sectional studies use calendars or other memory aids to collect data that describe the respondent’s recollection of an earlier time. Such data appear longitudinal and are quite useful. Memory of family characteristics in high school and attitudes, beliefs, and behaviors is more accurate when collected prospectively than when collected many years after the event (or perception of the event).³ Examples of factors that influence such retrospective memory are recall delay and telescoping. *Recall delay* refers to reductions in the report of behavior due to the respondent’s ability to remember events; the concept describes lower reports of remote events, characteristics, and behaviors than for more recent events. *Telescoping* refers to the allocation of events, characteristics, or behaviors to an earlier or later time period than the one in which it actually occurred.

Longitudinal data can inform researchers and policymakers about changes over time. Such data contain variables that allow better description and prediction of change, since these variables are based on current perceptions and not long-term recall. Possible descriptive questions that could be answered with longitudinal data include: What percentage of marijuana users continue to use marijuana in future years? How well can the movement of individuals between drug use and nonuse be described? Analysts of drug policy data have recommended a longitudinal study to describe and explain changes in drug use and criminal behavior over time.⁴ In addition to their usefulness in describing changes over time, longitudinal data can predict future behavior, using variables for prior behavior/attitudes without the weaknesses of retrospective memory.

¹ R. Simeone, W. Rhodes, D. Hunt, L. Truitt, *A Plan for Estimating the Number of “Hardcore” Drug Users in the United States: Preliminary Findings*, Abt Associates, 1997.

² W. Rhodes et al., *What America’s Users Spend on Illegal Drugs 1988-2000*, Office of National Drug Control Policy, Office of Programs, Budget, Research, and Evaluation, Executive Office of the President, 2001.

³ R. Tourangeau, L. J. Rips, and K. Rasinski, *The Psychology of Survey Response*, New York: Cambridge University Press, 2000; N Bradburn, L.J. Rips, S. K. Shevell, “Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys,” *Science*, v. 236, pp. 157-161, 1987.

⁴ *Informing America’s Policy on Illegal Drugs: What We Do not Know Keeps Hurting Us*, Committee on Data and Research for Policy on Illegal Drugs, Charles F. Manski, John V. Pepper, and Carol V. Petrie, Editors, Committee on Law and Justice and Committee on National Statistics, National Research Council, 2001.

The purpose of this study is to:

- Describe the movement of a representative sample of adolescents and young adults (14–21) into and out of drug use;
- Predict the correlates of “heavy drug use,” using factors of adolescence/early adulthood; and
- Present a preliminary method for screening and weighting a future sample of drug users.

Organization of this Report

This chapter provides the background and purpose of this report. Chapter 2 introduces the National Longitudinal Survey on Youth (NLSY) and the analytical techniques used in this study. Chapter 3 describes the movement into and out of drug use by adolescents and young adults between 1984 through 1998. Chapter 4 presents an analysis predicting heavy cocaine use. Chapter 5 presents a discussion on longitudinal screening.

2. METHODOLOGY

Description of National Longitudinal Survey of Youth

To describe drug use and to predict who becomes a heavy drug user, this study uses the Department of Labor's National Longitudinal Survey of Youth–1979 (NLSY79) cohort, representing 33.6 million youth. The NLSY79 is sponsored by the Bureau of Labor Statistics, Department of Labor, to focus on the labor market experiences of adolescents moving into the labor market and adulthood in the United States. For the last 23 years, the NLSY79 has followed a representative sample of about 10,000 youth (aged 14–21 as of December 31, 1978), through their late thirties and early forties by 2000. The survey oversamples blacks and Hispanics, permitting generalizations about these populations. Participants were interviewed annually until 1994 and biennially after 1994.

The NLSY79 has gathered information about the following factors that might influence the entry of adolescents/youth into the labor market:

- Alcohol and substance use;
- Attitudes and aspirations;
- Child care;
- Criminal behavior;
- Family background and demographic characteristics;
- Geographic residence;
- Health conditions, injuries, and insurance coverage;
- Household composition;
- Income and assets;
- Marital and fertility history;
- Military experience;
- Schooling, school records, and aptitude information;
- Training investments; and
- Work and non-work experiences.

The NLSY began in 1979 with a sample of 12,686 youth and continues to have a high response rate (see Exhibit 2.1). In 1980, 12,141 individuals were interviewed with a response rate of 96 percent. By 1990, the response rate was 91 percent. In 1998, the last year included in this study, the response rate was 87 percent, after 17 rounds and 19 years. By 2000, the age of the sample population ranged from 35–42.

This study uses the battery of questions on drug use that were administered in 1984, 1988, 1992, 1994, and 1998. Comparable questions were asked regarding lifetime and past-year use of marijuana, cocaine (between 1984 and 1998), and crack (between 1992 and 1998).

Exhibit 2.1: NLSY79 Mode and Response Rates, by Year

YEAR	MODE	SAMPLE SIZE	RESPONSE RATE
1979	Personal/PAPI	12,686	—
1980	Personal/PAPI	12,141	96%
1981	Personal/PAPI	12,195	96%
1982	Personal/PAPI	12,123	96%
1983	Personal/PAPI	12,221	97%
1984	Personal/PAPI	12,069	96%
1985	Personal/PAPI	10,894 ⁵	95%
1986	Personal/PAPI	10,655	93%
1987	Telephone/PAPI	10,485	91%
1988	Personal/PAPI	10,465	91%
1989	Personal/PAPI /CAPI	10,605	93%
1990	Personal/PAPI/CAPI	10,436	91%
1991	Personal/PAPI	9,018 ⁶	92%
1992	Personal/PAPI	9,015	92%
1993	Personal/CAPI	9,011	92%
1994	Personal/CAPI	8,891	91%
1996	Personal/CAPI	8,636	89%
1998	Personal/CAPI	8,399	87%
2000	Personal/CAPI	8,033	83%

SOURCE: U.S. Department of Labor, Bureau of Labor Statistics, *NLS Handbook, 2000 and 2001*.

⁵ A total of 201 military respondents were retained from the original sample of 1,280; 186 of the 201 participated in the 1985 interview. The total number of NLSY79 civilian and military respondents eligible for interview (including deceased respondents) beginning in 1985 was 11,607.

⁶ The 1,643 economically disadvantaged non-black/non-Hispanic male and female members of the supplemental subsample were not eligible for interview in the 1991 survey year. The total number of NLSY civilian and military respondents eligible for interview (including deceased respondents) beginning in 1991 was 9,964.

Analytic Techniques

This study focuses on two general areas of interest: (1) the movement of individuals into and out of drug use and (2) the prediction of heavy drug use. Each will be presented in turn.

DESCRIPTION OF MOVEMENT INTO AND OUT OF DRUG USE

In order to depict movement into and out of drug use, we developed “distribution trees” for the five index years during the 15-year period between 1984 and 1998. The distribution trees are based on frequency counts of use starting with 1984 and conditional frequency counts for subsequent years. Frequency counts for 1988 are based on the condition of use or nonuse for 1984, and frequency counts for 1992 are based on condition of use or nonuse for the 1984 and 1988 survey years. Frequency counts for 1994 are in turn based on condition of use or nonuse for the earlier index years. For 1998, the counts are based on condition of use or nonuse for all previous study periods. For example, the distribution of those who used in 1988 is based on use versus nonuse in 1984. For marijuana, cocaine, and a combination of both, two distribution trees were developed—one for those who reported use and another for those who did not report use, starting with the 1984 period.

The trees are also based on crosstabulations (2*2 tables) that yielded distributions of use and nonuse for two adjacent study periods. Crosstabulations were used to verify the conditional frequency counts developed for the distribution trees. One additional advantage is that these trees permit recalculation of change by combining each of the component cells.

PREDICTION

The prediction of heavy drug use primarily employs logistic regressions and odds ratios. No claim is made for a direct causal relationship among the variables used in the analysis. Rather, we seek to explain whether certain behaviors or characteristics tend to coincide with the presence of heavy drug use to a greater or lesser degree than do other variables or characteristics.

Logistic regression is often used when the dependent variable is composed of two values, such as the presence and absence of an activity (e.g., was a heavy drug user or was not a heavy drug user). The measure of strength of association is r^2 .

To illustrate the model, let H denote the measurement of the dichotomous outcome, heavy drug use. Then $H=0$ if the individual was not a heavy drug user and $H=1$ if the individual was a heavy drug user. Variables W through Z symbolize additional explanatory variables, which may be either dichotomous or continuous. (NOTE: The three explanatory variables are presented for illustrative purposes only; additional explanatory variables do not change the form of the model.). The unified model for dichotomous outcomes is:

$$\text{Logit}(H) = b_0 + b_1*W + b_2*X + b_3*Z$$

Logistic regression generates odds ratio estimates for each predictor. Such estimates are readily interpretable probabilities that indicate how much more likely it is that an outcome would be observed if, all other elements being the same, the predictor occurs compared to when the predictor

does not occur. For example, all other things being equal, an odds ratio would estimate how much more or less likely a man than a woman is to become a heavy drug user. An odds ratio above 1.0 means that the activity becomes more likely, while an odds ratio below 1.0 means that this activity becomes less likely. For continuous variables, we used ordinary least-squares models.

Variables

Chapter 3 describes drug use for five index years between 1984 and 1998. To portray the movement into and out of drug use, the variables presented are marijuana use, cocaine use, and both marijuana and cocaine use in the past year.

Chapter 4 focuses on predicting drug use. The dependent variable, heavy drug use, is the first focus. There are many ways that the concepts and measures related to heavy drug use have been operationally defined. One operational measure that the Substance Abuse and Mental Health Services Administration (SAMHSA) used in its National Household Survey on Drug Abuse (NHSDA) to define and count individuals who are dependent on drugs is *problematic drug use*. This measure is based on criteria established by the American Psychiatric Association in its *Diagnostic and Statistical Manual of Mental Disorders* versions three (revised) and four (DSM-III-R, DSM-IV). From 1994 to 2000, SAMHSA applied the DSM-IV definition that a person was dependent if s/he met three out of seven dependence criteria⁷ (for substances with a withdrawal criterion) or three out of six criteria (for substances without a withdrawal criterion) based on responses to questions in the NHSDA. These questions were revised for the 2000 survey, and questions were added, regarding respondents' potential abuse of drugs (again, based on DSM-IV criteria).

Chronic drug use is a concept used in ONDCP's publication, *What America's Users Spend on Illegal Drugs* (2001). The study defined chronic users by drug use behaviors rather than DSM criteria. The report's authors analyzed the data collected by the National Institute of Justice's Drug Use Forecasting program and operationally defined the concept of chronic drug use as "those who admitted using cocaine or heroin on more than 10 days during the month before being arrested."⁸

Obviously, the differences among these measures of heavy drug use are substantial, ranging from a combination of psychological variables (i.e., the desire for drugs, the difficulty of stopping use, and withdrawal when attempting to end use) as opposed to measures of actual use. The NLSY questions on illicit drugs focus on behaviors regarding drug use, rather than on dependence or abuse. In addition, the categories used in the NLSY to measure level of drug use (e.g., marijuana, cocaine, crack) are high, but capped at 100 or more lifetime uses.⁹ That volume of drug use is substantial, but does not contain information on attempts at withdrawal from drug use or on the difficulty of withdrawal.

⁷ The operational definitions before and after 2000 are described in: Substance Abuse and Mental Health Administration, *National and State Estimates of the Drug Abuse Treatment Gap: 2000 National Household Survey on Drug Abuse*, Appendix A, DHHS, 2002.

⁸ W. Rhodes, *ibid.*

⁹ There is one exception. In 1984, the measure was capped at 1,000 or more uses.

Eighteen rounds of the NLSY were conducted between 1979 and 1998.¹⁰ Questions on drug use were administered for five of these rounds between 1984 and 1998. Based on data from these years, we developed an operational definition for heavy drug use based on the following criteria:

- Include drugs generally associated with very heavy drug use (cocaine, crack, heroin) as feasible; and
- Select operational definitions that provide a robust number of cases for analysis.

The operational definition of heavy drug use for the analysis in Chapter 4 is both (1) self-reported lifetime use of cocaine or crack on 100 or more occasions, and (2) cocaine or crack use in the year of the interview (it is important to note that not all individuals in the sample responded each year). Heavy cocaine users in any of the five survey years were retained, even if they were nonrespondents to any of the first four rounds (e.g., a “heavy” cocaine user during 1984 was kept in the analyses even if s/he was a nonrespondent after 1984). As presented in Exhibit 2.2, a lenient measure—report of lifetime cocaine or crack use on 100 or more occasions—results in 1,447 heavy drug users. The more stringent measure (e.g., 100 or more lifetime uses of cocaine or crack plus use on 3–4 days per week of cocaine/crack) results in only 127 heavy drug users. We selected a measure that provides enough cases for a robust analysis and also contains a population that used a substantial amount of drugs. This measure—self-reported use of cocaine or crack 100 or more times in one’s lifetime plus cocaine or crack use in the past year— results in 551 individuals defined as heavy drug users.

Other drugs were not included in the operational definition of heavy drug use for several reasons. Heroin and injection drug use were excluded, as the NLSY asked only whether or not the respondent had ever injected drugs. Such knowledge about use without indication of frequency of use did not seem *a priori* heavy drug use. In addition, lifetime use of marijuana was not included as a measure of lifetime heavy drug use, as the operational definition (100 or more times) seemed too low a standard. (NOTE: The range of individuals with 100 or more lifetime uses of marijuana ranges from 2,093 in 1984, to 1,141 in 1992, and 1,330 in 1998. The numbers of individuals who used marijuana in the year prior to each NLSY round are as follows: 3,812 in 1984, 2,213 in 1988, 1,084 in 1992, 1,173 in 1994, and 814 in 1998).

¹⁰ The NLSY was also conducted in 2000 and 2002, and another round will be conducted in 2004; none included questions on drug use.

Exhibit 2.2: Drug Use for Individual Years and Combinations of Year

	1984	1988	1992	1994	1998	ANY of the 5
Cocaine—Lifetime use 100+ times	240	273	583	315	294	1,335
Cocaine—Used in last year	302	990	323	302	184	2,085
Cocaine—Lifetime use 100+ and used cocaine in past year	191	165	39	93	72	484
Cocaine—Lifetime use 100+ and used cocaine 1–2 days or more per week	82	32	8	26	16	156
Cocaine—Lifetime use 100+ and used cocaine 3–4 days or more per week.	50	20	8	16	12	103
Cocaine or crack—Lifetime use 100+ and used cocaine or crack in past year	191*	165*	71	118	106	551
Cocaine or crack—Lifetime use 100+ and used cocaine/crack 1–2 days or more per week	82*	32*	17	43	29	189
Cocaine or crack—Lifetime use 100+ and used cocaine/crack 3–4 days or more per week	50*	20*	16	24	22	127
Heroin—Lifetime use	119	n/a	126	199	170	416
Injection drug use (without doctor's permission)—Lifetime use	n/a	117**	n/a	206	198	382
Cocaine or crack, Lifetime 100+	240*	273*	678	361	349	1,447
Cocaine or crack 100+ or heroin	329	273*	785	462	430	1,613
Cocaine or crack 100+ or heroin or injection drug use	329	340	785	544	500	1,723
Cocaine or crack, lifetime use 100+, heroin or IV drug use, and used cocaine/crack in past year	222	165	103	153	115	615
Cocaine or crack, Lifetime use 100+ Heroin or IV drug use, and used cocaine/crack 1–2 days or more per week	84	32	32	48	30	209
Cocaine or crack, Lifetime use 100+, Heroin or IV drug use, and used cocaine/crack 3–4 days or more per week	51	20	30	28	23	145

* Powder cocaine only; **Cocaine injectors

As cocaine and crack are the operational measures, we refined the dependent variable to heavy cocaine use. Cocaine and crack were extensively used in the 1980s and 1990s, the period analyzed in this study.

Nonresponse Analysis

As with any longitudinal survey, dropouts occur over time. In a study concerned with drug use, it is reasonable to determine if heavy cocaine users (defined above) were more likely to drop out. If, for example, heavy cocaine users were more likely to drop out, the data set may not be representative of all heavy cocaine users who were 14–21 years old on December 31, 1978.

Fortunately, we can test the hypothesis of whether heavy cocaine users were more likely to drop out of the NLSY79 than other youth, using the method of Diggle.¹¹ Since we have five study periods (1984, 1988, 1992, 1994, and 1998) with data on drug use, this method calls for a separate test on each of the first four time periods.

The first test is whether those who drop out after 1984 (they completed the 1984 interview, but they were not respondents in the 1988, 1992, 1994, or 1998 interviews) are more likely to be heavy cocaine users than those who are respondents in 1984. Of the 12,069 respondents in 1984, 10,700 responded in at least one of 1988, 1992, 1994, and 1998, while 1,369 were nonrespondents to all four as shown in Exhibit 2.3. Among the 10,700 non-dropouts, 1.57 percent satisfied the heavy cocaine user condition in 1984 (100 or more lifetime uses of cocaine or crack, plus use of either cocaine or crack in the last year). It is important that we only use 1984 data to determine the heavy cocaine users, since we only have data beyond 1984 for the non-dropouts. The percentage of heavy cocaine users was 1.57 percent among the non-dropouts and 1.68 percent for the dropouts; this difference was not statistically significant ($p=.76$).

The second test for those who responded in 1988 is very similar. We split this subset into dropouts (nonrespondents in 1992, 1994, and 1998) and nondropouts, and determined the percentages in each who satisfied the heavy cocaine use condition in 1984 or 1988. This test also returned a non-significant p value ($p=.43$). The third (1992) and fourth (1994) tests were conducted in similar fashion, and the differences were not statistically significant. No test is appropriate for 1998, of course, since this is the most recent year drug-related questions were asked in the NLS79.

In conclusion, NLSY79 dropouts were *not* more likely to have been heavy cocaine users before dropping out of the survey than were non-dropouts.

¹¹ P. J. Diggle, K. Y. Liang, and S. L. Zeger, *The Analysis of Longitudinal Data*, Oxford, England: Oxford University Press, 1994.

Exhibit 2.3: Four Tests of Non-Random Dropout among Heavy Cocaine Users

		Dropout?	
		Yes	No
1984	Number of cases	1,369	10,700
	Percent heavy cocaine users	1.68%	1.57%
	Chi-square statistic	0.0942	1 d.f.
	<i>p</i> value	0.76	Not significant
1988	Number of cases	1,704	8,761
	Percent heavy cocaine users	3.11%	2.76%
	Chi-square statistic	0.6309	1 d.f.
	<i>p</i> value	0.43	Not significant
1992	Number of cases	216	8,800
	Percent heavy cocaine users	2.31%	3.41%
	Chi-square statistic	0.7724	1 d.f.
	<i>p</i> value	0.38	Not significant
1994	Number of cases	695	8,196
	Percent heavy cocaine users	5.32%	4.34%
	Chi-square statistic	1.4569	1 d.f.
	<i>p</i> value	0.23	Not significant

3. DRUG USE 1984–1998

This section compares the sample analyzed with the sample not included in the analysis, describes the movement in and out of drug use, and then compares the NLSY with another national survey, the National Household Survey on Drug Abuse (NHSDA).

Comparison Between Those Included/Not Included in Analysis

In Exhibits 3.1–3.8, we present tables that depict use and nonuse of illicit drugs by individuals between 1984 and 1998. In order to determine movement in and out of drug use, we selected individuals who had consistently responded in survey years 1984, 1988, 1992, 1994, and 1998 of the NLSY. By selecting individuals who had participated in each round of the survey, we avoided developing a database with gaps due to nonresponse in some of the survey years. The analyses presented in this chapter are based on the individuals who responded to all five surveys with drug use modules between 1984–1998, and some who responded to every year except 1998 (participation in 1998 was not a condition for inclusion in the analysis)¹². In all, 7,597 individuals responded to all five rounds, and an additional 534 responded to only the first four rounds (8,131 in total).

In order to examine the effect of this decision, we conducted two analyses in order to learn the following: (1) Are the individuals who did not respond in 1998, but responded in 1984, 1988, 1992 and 1994 (n=534), different from those who responded every year including 1998 (n=7,597)? (2) Is the sample included in the analysis (7,597+534) different from those we excluded because they did not respond to the NLSY in one or more years between 1984 and 1994?

With regard to question 1, Exhibit 3.1 presents a comparison along 17 variables between respondents in 1984–1998 and respondents in 1984–1994. The results indicate few differences between the two groups. More females were in the 1984–1998 group, and there was a significantly higher percentage of whites in the 1984–1998 group (79.2 percent vs. 75.7 percent). Significantly more individuals in the 1984–1994 group had used marijuana or hashish in 1984, had used cocaine in survey years 1984, 1992, and 1994, and had been suspended from school. For 12 variables, these two groups were not statistically different: There were no significant differences for Hispanic and black race/ethnicity; high school education; and whether the respondent sold hard drugs, gained income from illegal activities in the last year, smoked cigarettes in the past 30 days, smoked marijuana or hashish in 1988–1994, and used cocaine in 1988. Results suggest that these two groups do not differ markedly, but enough that it is desirable to include the 1984–1994 respondents in our analyses.

Exhibit 3.2 compares the sample used for our study with those we excluded. The study sample includes the group who consistently responded to all five rounds (1984–1998) and four rounds (1984–1994); this group is labeled the “in study group.” Nonrespondents (or those excluded from the sample) are those who did not respond in one or more years between 1984 and 1994; this group is labeled “out of study.” As in the previous analysis, these groups were compared along 17 characteristics.

¹² 1998 is the last survey year used in the analysis and nonresponse in that year would not affect comparisons to subsequent survey years.

Exhibit 3.1: Comparisons of Sample Respondents 1984–1998 And Sample Respondents 1984–1994

	Respondents To All 5 Rounds: 1984–1998 (Base <i>n</i> max=7,597)		Respondents To 4 Rounds: 1984–1994 (Base <i>n</i> max=534)	
	%	N	%	N
Percentages				
Female*	51.9	7,597	38.9	534
Hispanic	6.3	7,597	7.5	534
White (non-Hispanic)*	79.2	7,597	75.7	534
Black (non-Hispanic)	14.5	7,597	16.9	534
At least high school education	38.9	7,597	36.6	534
Smoked marijuana or hashish in 1984*	33.1	7,597	40.0	534
Smoked marijuana or hashish in 1988	22.7	7,597	22.6	534
Smoked marijuana or hashish in 1992	13.7	7,597	14.6	534
Smoked marijuana or hashish in 1994	14.3	7,597	14.3	534
Used cocaine in 1984*	11.4	7,597	14.2	534
Used cocaine in 1988	10.6	7,597	10.4	534
Used cocaine in 1992*	3.6	7,597	5.2	534
Used cocaine in 1994*	3.0	7,597	5.5	534
Ever been suspended from school (1980)*	20.8	7,404	25.1	516
Sold hard drugs in past year (1980)	2.4	7,201	3.0	500
Illegal income in past year (1980)	16.1	7,032	16.9	476
Smoked cigarettes in past 30 days (1984)	40.7	7,570	42.6	533
Note: Significant differences between percentages were based on chi-square tests of independence.				
* Difference between respondents and nonrespondents is significant at $p < 0.05$.				

Results show that these groups differ significantly for eight characteristics. Hispanics were more likely to be in the out of study group, and individuals in the in-study group were more likely to be female. Individuals in the out of study group were more likely to have a high school education (42.5 percent vs. 38.7 percent).

Exhibit 3.2 Comparisons of Study Sample and Sample Excluded from Study

	In-Study Group (Base <i>n</i> max=8,131)		Out of Study Group (Base <i>n</i> max=1,855)	
	%	<i>n</i>	%	<i>n</i>
Percentages				
Female*	51.1	8,131	43.4	1,855
Hispanic*	6.3	8,131	8.3	1,855
White (non-Hispanic)	79.0	8,131	77.6	1,855
Black (non-Hispanic)	14.7	8,131	14.1	1,855
At least high school education*	38.7	8,131	42.5	1,855
Smoked marijuana or hashish in 1984*	33.5	8,131	30.4	1,376
Smoked marijuana or hashish in 1988*	22.7	8,131	17.9	848
Smoked marijuana or hashish in 1992*	13.8	8,131	10.3	885
Smoked marijuana or hashish in 1994*	14.3	8,131	10.8	760
Used cocaine in 1984*	11.6	8,131	9.3	1,376
Used cocaine in 1988	10.5	8,131	8.7	848
Used cocaine in 1992	3.7	8,131	3.4	885
Used cocaine in 1994	3.2	8,131	3.2	760
Ever been suspended from school (1980)	21.1	7,920	22.0	1,650
Sold hard drugs in past year (1980)	2.4	7,701	2.5	1,562
Illegal income in past year (1980)	16.1	7,508	14.7	1,522
Smoked cigarettes in past 30 days (1984)	40.8	8,103	39.8	1,368
Note: Significant differences between percentages were based on chi-square tests of independence.				
* Difference between respondents and nonrespondents is significant at $p < 0.05$.				

Respondents in the study were more likely than those out of the study to have smoked marijuana in 1984, 1988, 1992, and 1994, and to have used cocaine in 1984. The group included in this analysis had higher rates of drug use than those excluded from the analysis, leaving us less concerned about the possible diminution of drug users from the analysis.

These results show that the respondents included in the study share more in common with respondents excluded from the study; that is, these two groups are more alike than dissimilar, especially because there were no significant differences between them for a majority of the variables along which they were compared. Although there are notable differences, particularly for the behavioral variables regarding drug use, these differences are in the right direction; for instance, there are higher proportions of individuals in the in-study group who have smoked marijuana, which is useful for this study. The usual concern is that those excluded from the analysis are less likely to engage in the analyzed behavior than those included. For this analysis of the movement into and out

of drug use, those included are somewhat more likely than those in the out-of-study group to have smoked marijuana. Otherwise, the differences between the groups are small.

Movement Into and Out of Drug Use

MARIJUANA USE IN THE PAST YEAR

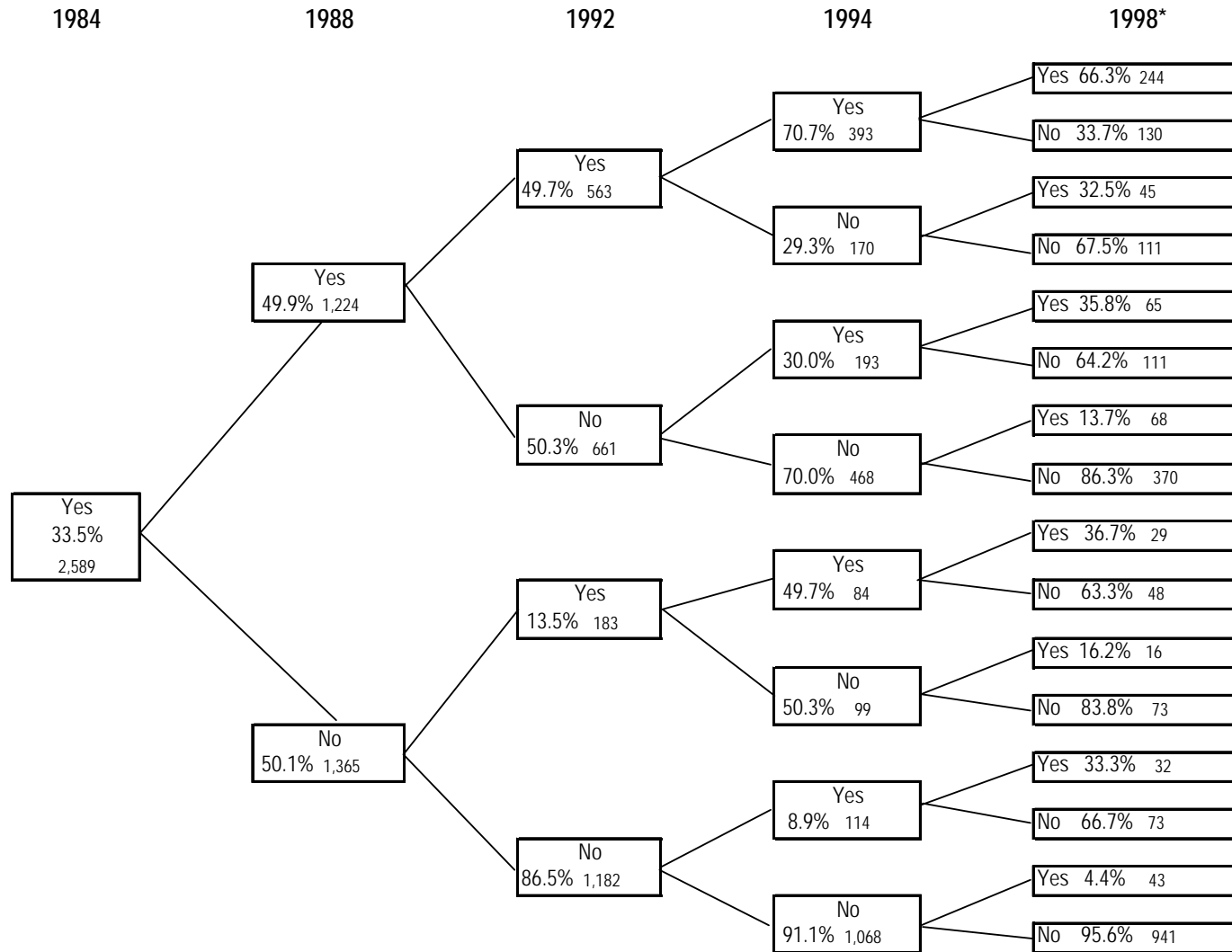
Exhibit 3.3 depicts the distribution “tree” of respondents who reported marijuana use in the past year, in 1984; Exhibit 3.4 depicts the distribution tree of respondents who stated they did not use marijuana in 1984.

The tree describes the distribution of individuals who reported use or nonuse of drugs in each of the five index years from 1984 to 1998. There are two distribution trees for each drug. The first is based on respondents who reported marijuana use in 1984, and the second, on respondents who reported no marijuana use in 1984. Subsequent years after 1984 describe the distributions of respondents based on whether they reported marijuana use or no marijuana use. For instance, the 1988 column in Exhibit 3.3 describes the distribution of respondents who reported marijuana use (“yes”) or no marijuana use in that year (“no”) after reporting drug use in 1984. In 1992, the distribution of respondents is expanded to capture drug use or non-use based on the distribution in 1988. In 1992, there are two binomial distributions (i.e. four alternatives); one is based on those who reported drug use and the other based on those who did not report drug use in 1988. In 1994, there are four binomial distributions based on responses in 1992. And in 1998, there are eight binomial distributions based on responses in 1994. In this way, the distribution tree presents movement in and out of drug use for the five index years (1984, 1988, 1992, 1994, and 1998) over the 15-year study period based on the initial response to questions on drug use in 1984. Weighted percentages are used in all the exhibits.

Over the first 5-year timespan (1984–1988), half of all individuals who used marijuana in 1984 did not use again in 1988 (see Exhibit 3.3). Over the next 5-year span (1988–1992), almost half of those who had used in 1988 reported use again in 1992, and almost 87 percent of those who did not use in 1988 did not use again in 1992. From 1992 to 1998, one-half to two-thirds of individuals who used in one index year also used in the subsequent index year, if they initiated use in 1984 or 1988. Overall, only 3 percent (244/8,131) used marijuana in every one of the five interview years.

Exhibit 3.4 shows that two-thirds (66.5 percent) of respondents did not use marijuana in 1984, and 91 percent (5,062/5,542) did not report use in the subsequent study year. Of those who used in 1988, three-quarters reported no use in 1992. Overall, almost all those who did not use marijuana in 1984 did not use in 1998 (96 percent, or 4,990/5,198). These results suggest that youth who do not initiate marijuana use in the early years are unlikely to ever start marijuana use.

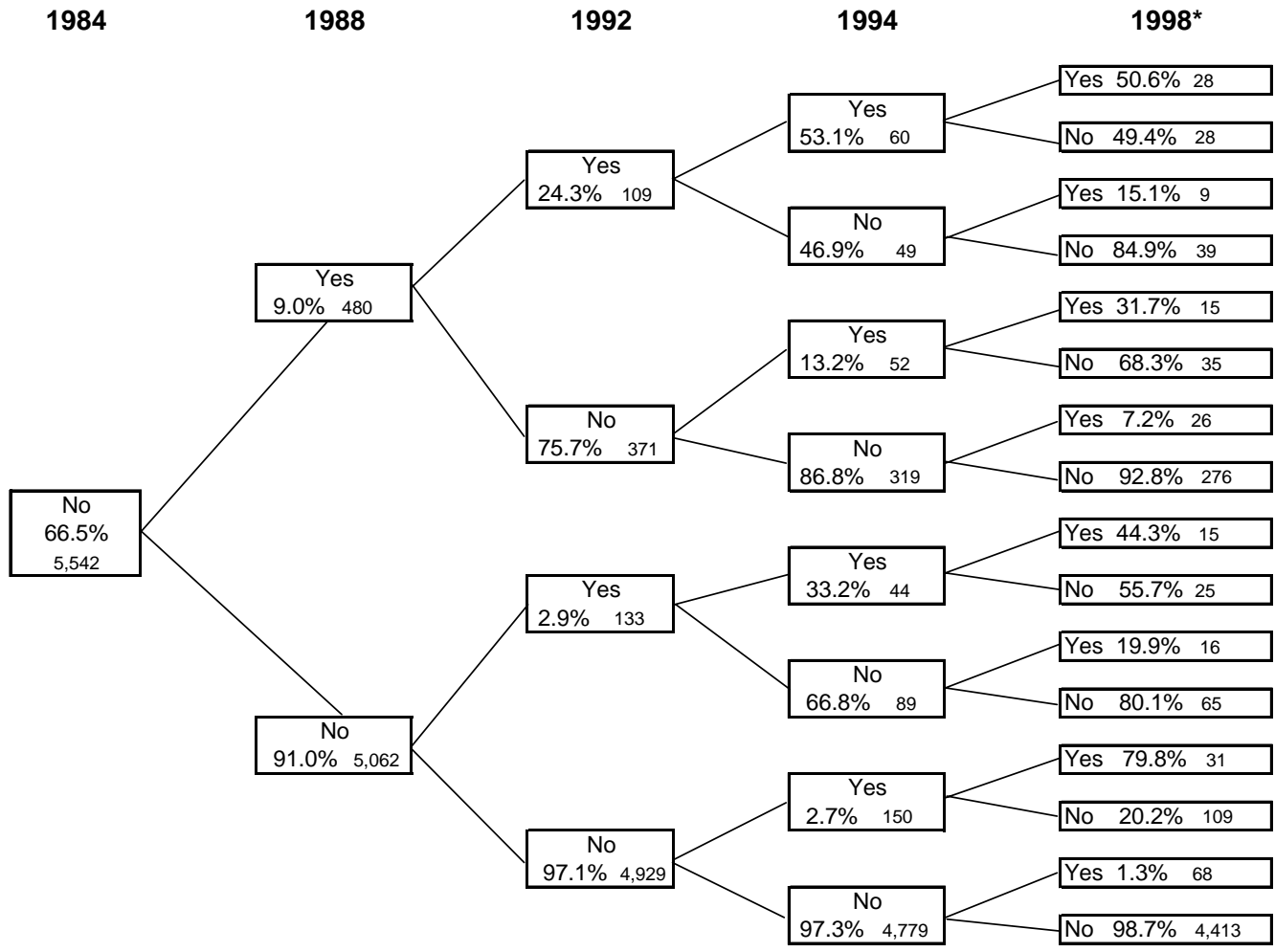
Exhibit 3.3: Distribution Tree of Respondents Who Reported Smoking Marijuana in 1984



Source: National Longitudinal Study of Youth, 1979 cohort, 1984-1998

* There were 534 individuals who responded to the 1984, 1988, 1992, and 1994 surveys, but not to 1998. These individuals were included in 1984-1994, but not in 1998.

Exhibit 3.4: Distribution Tree of Respondents Who Reported No Marijuana Use in 1984



Source: National Longitudinal Study of Youth, 1979 cohort, 1984-1998

* There were 534 individuals who responded to the 1984, 1988, 1992, and 1994 surveys, but not to 1998. These individuals were included in 1984-1994, but not in 1998.

These results also suggest that few adolescents who initiate marijuana use continue such behavior consistently over the long run. In general, most of those who did not use marijuana in 1984 did not use over the five index years during the 15-year period. However, many adolescents also reported intermittent marijuana use over the 15-year period.

COCAINE USE IN THE PAST YEAR

Exhibits 3.5 and 3.6 present data on cocaine use during the interview years of 1984 to 1998. Whereas cocaine users are more prevalent than marijuana users, there are similar patterns of use for marijuana and cocaine.

When cocaine use is examined over the first 5-year timespan (1984–1988), we see that two-fifths who initiated cocaine use in 1984 used again in 1988. Over the next five year span (1988–1992), a quarter of those who had used in 1984 and 1988 used again in 1992, and only 1 in 10 individuals who had not used in 1988 used in 1992. Overall, less than two-tenths of one percent used cocaine in every year over the study period, and of those who reported cocaine use in 1984, 93 percent (692/748) did not use in 1998.

Exhibit 3.6 shows that about seven-eighths (88.4 percent) of interviewees did not report cocaine use in 1984, and over nine-tenths of these individuals (93.7 percent) did not use in 1988, the subsequent study year. Of the few (6.7 percent) individuals who used cocaine in 1988, more than four-fifths (85.6 percent) did not use in 1992. Almost all respondents who did not use cocaine in 1984 did not use in 1998 (98 percent, or 6,737/6849).

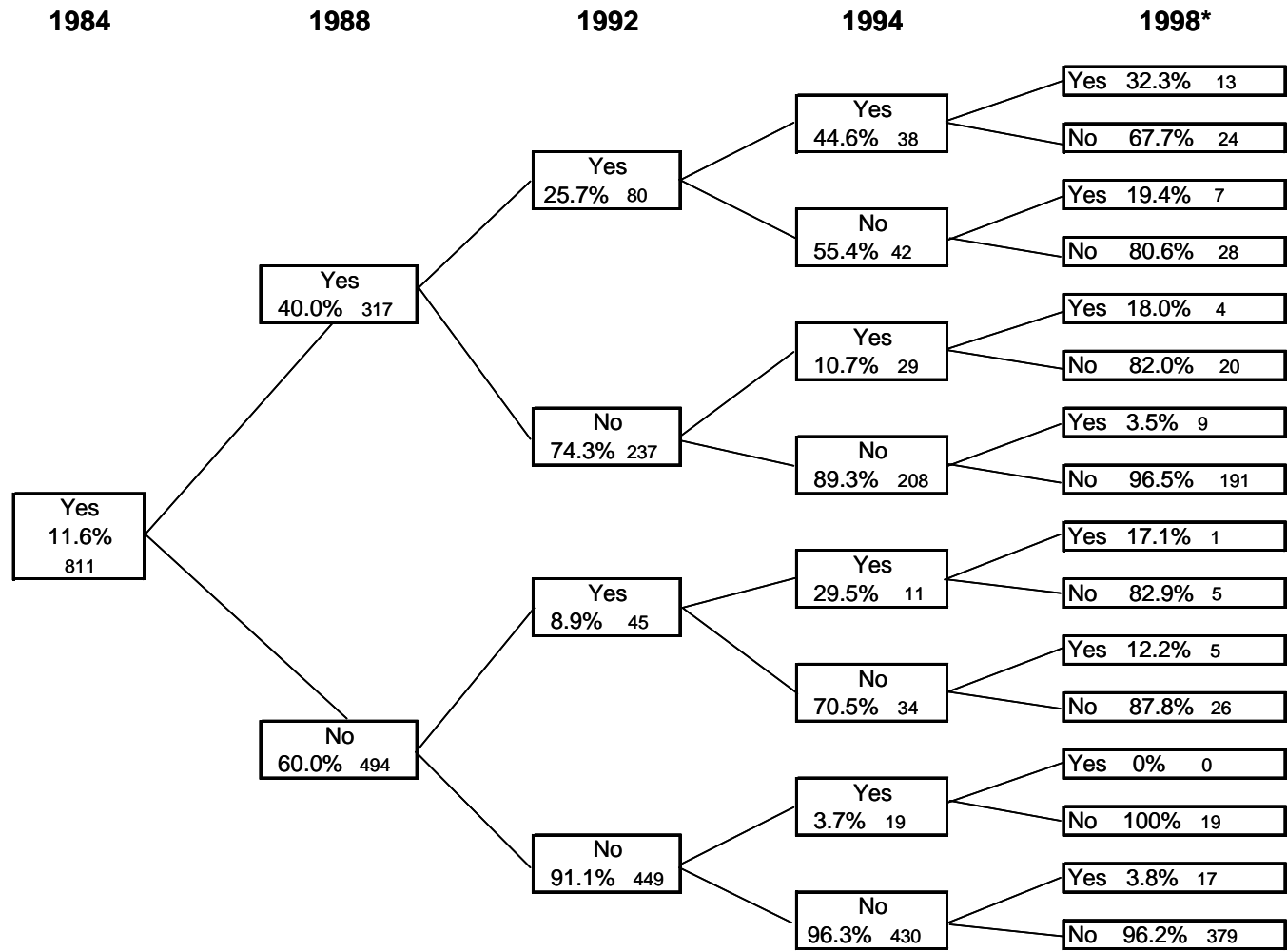
MARIJUANA AND COCAINE USE IN THE PAST YEAR

Ten percent of respondents reported use of both marijuana and cocaine in 1984 (see Exhibit 3.7). Of these individuals, 35.7 percent used both drugs again in 1988 (3.6 percent of interviewees). In 1992, almost one quarter (23.8 percent) of those who used in 1988 and 1984 used both drugs again.

Exhibit 3.8 shows that nine-tenths of respondents (89.8 percent) did not use marijuana and cocaine in 1984. Only 5 percent of those who did not use both drugs in 1984 used both drugs in 1988. Of the 5 percent who reported use of marijuana and cocaine in 1988, approximately 9 out of 10 did not use both again in 1992. Overall, almost all non-users of marijuana and cocaine in 1984 did not report use in the subsequent interview years from 1988 to 1998 (98.9 percent, or 6,872/6,948). These data from combined marijuana and cocaine use paint a similar story to the results presented in Exhibits 3.3 to 3.6. Though most respondents do not report initiating drug use, there are others who initiate and use these drugs intermittently.

Even though the relationship between initiation of drug use and the continuation of such behavior is not clear from this analysis, the results clearly depict individuals moving in and out of drug use. In general, these data suggest that the majority of youth abstained from drug use during the study period, and only a few individuals consistently used drugs in every interview or index year (3.2 percent for marijuana and 0.2 percent for cocaine). Further, varying proportions of individuals who initiated drug use became intermittent users or regular users. What these data do not tell us are the intervening factors that influence movement in and out of drug use.

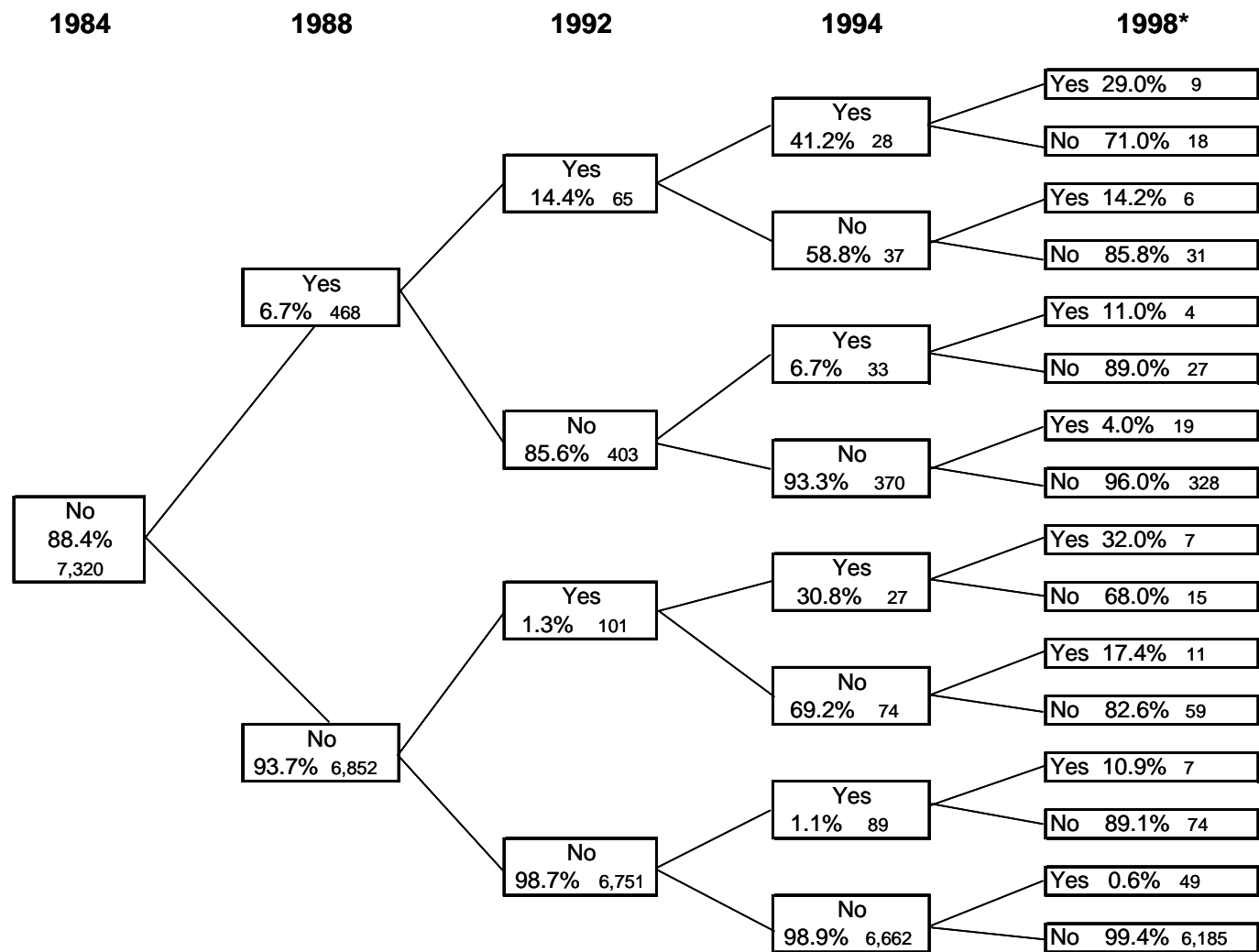
Exhibit 3.5: Distribution Tree of Respondents Who Reported Using Cocaine in 1984



Source: National Longitudinal Study of Youth, 1979 cohort, 1984-1998

* There were 534 individuals who responded to the 1984, 1988, 1992, and 1994 surveys, but not to 1998. These individuals were included in 1984-1994, but not in 1998.

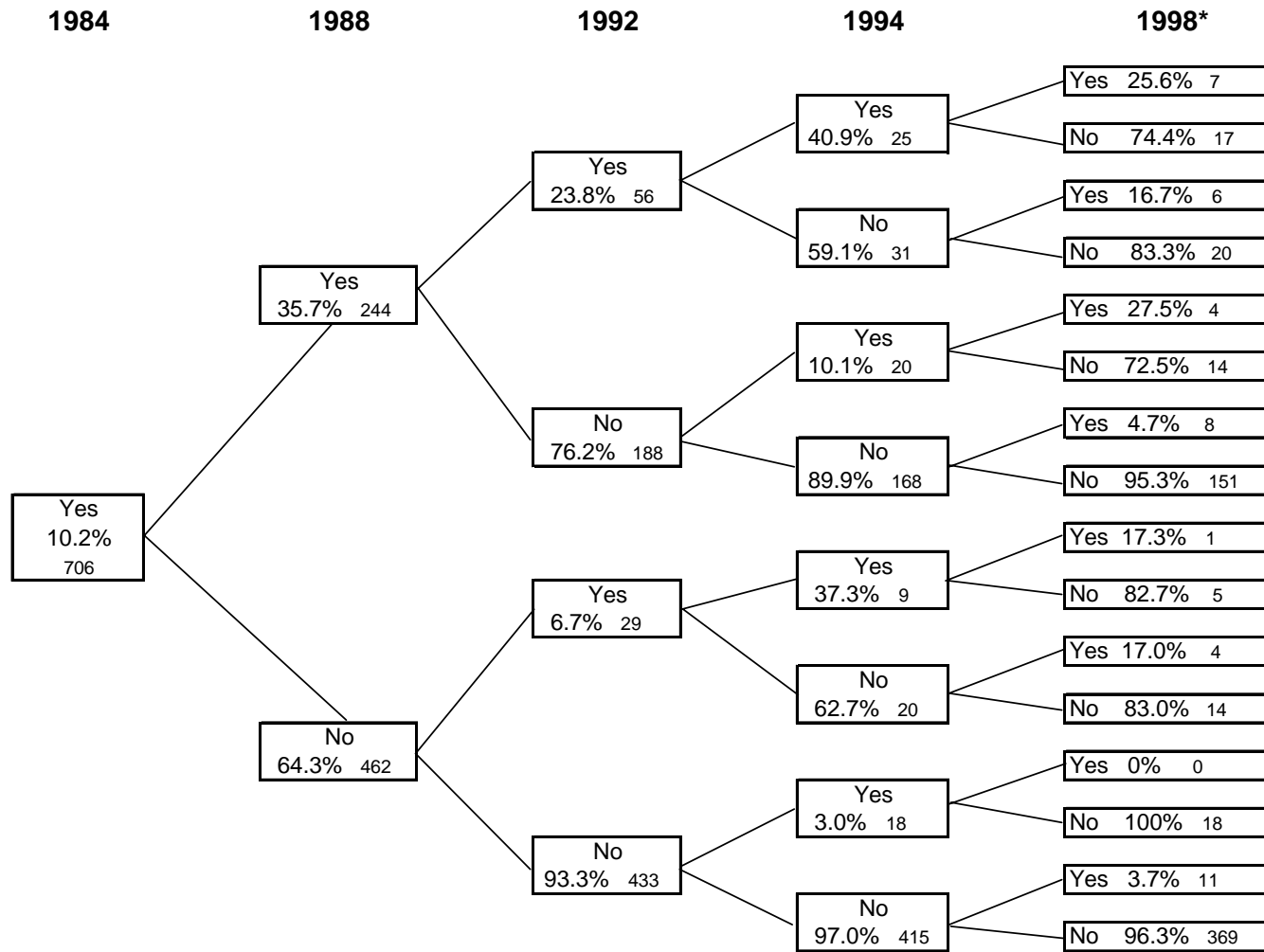
Exhibit 3.6: Distribution Tree of Respondents Who Reported No Cocaine Use in 1984



Source: National Longitudinal Study of Youth, 1979 cohort, 1984-1998

* There were 534 individuals who responded to the 1984, 1988, 1992, and 1994 surveys, but not to 1998. These individuals were included in 1984-1994, but not in 1998.

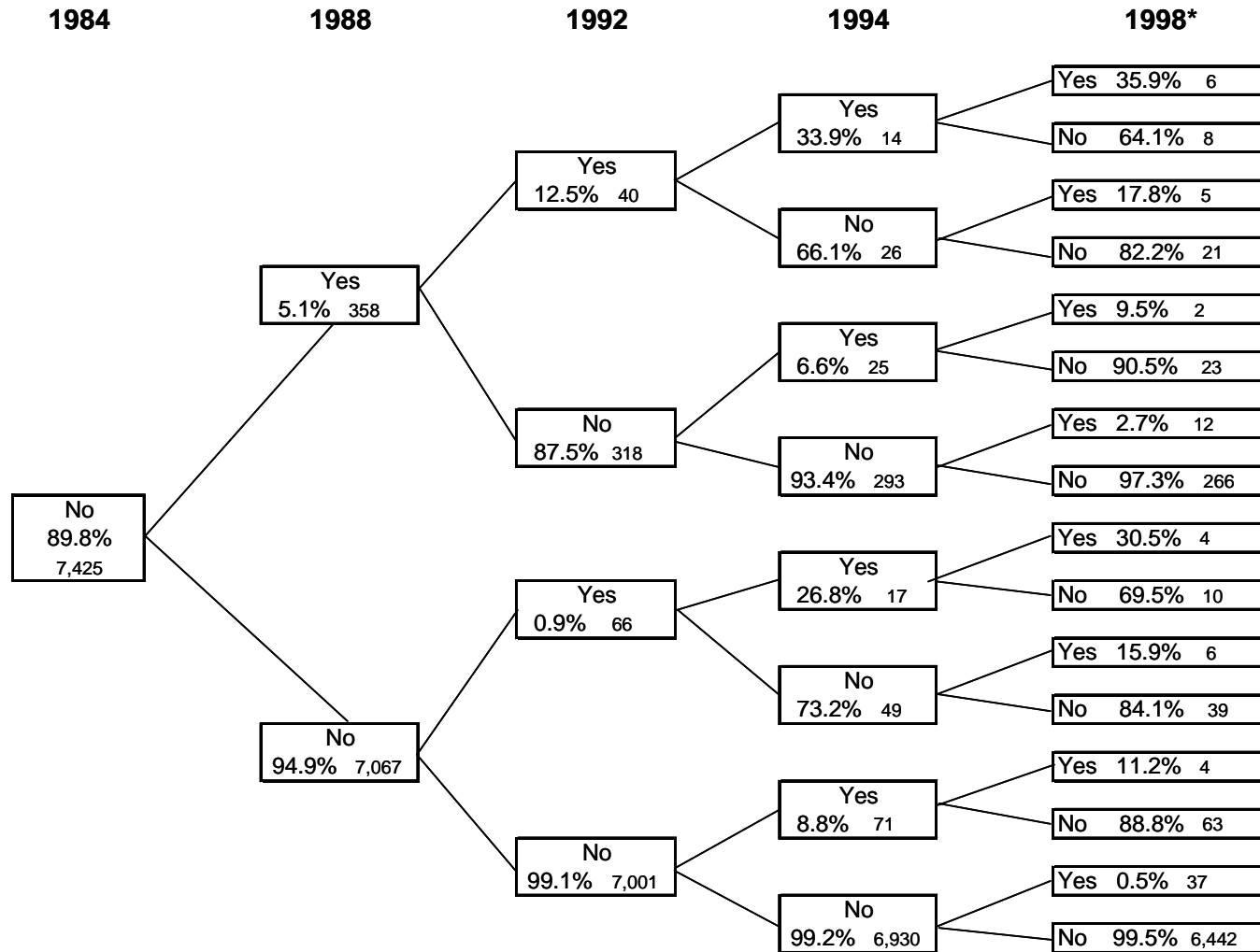
Exhibit 3.7: Distribution Tree of Respondents Who Reported Cocaine and Marijuana Use in 1984



Source: National Longitudinal Study of Youth, 1979 cohort, 1984-1998

* There were 534 individuals who responded to the 1984, 1988, 1992, and 1994 surveys, but not to 1998. These individuals were included in 1984-1994, but not in 1998.

Exhibit 3.8: Distribution Tree of Respondents Who Reported No Cocaine and Marijuana Use in 1984



Source: National Longitudinal Study of Youth, 1979 cohort, 1984-1998

* There were 534 individuals who responded to the 1984, 1988, 1992, and 1994 surveys, but not to 1998. These individuals were included in 1984-1994, but not in 1998.

CORRELATES OF NUMBER OF YEARS USED DRUGS: GENDER AND RACIAL/ETHNIC DIFFERENCES

Exhibits 3.9 and 3.10 display the association between demographic characteristics and the number of index years respondents used marijuana and cocaine. Higher percentages of females than males abstain from marijuana and cocaine use. For both marijuana and cocaine, males are about twice as likely as females to have used in 3–5 index years.

Exhibit 3.9: Number of Times Used Marijuana By Gender*

Number of Index Years	Male	Female
Used Marijuana		
0	49.4%	66.2%
1	21.4%	17.5%
2	11.8%	7.7%
3–5	17.4%	8.7%

* $p < 0.05$

Exhibit 3.10: Number of Times Used Cocaine By Gender

Number of Index Years	Male	Female
Used Cocaine		
0	75.7%	86.8%
1	14.9%	8.8%
2	6.3%	3.1%
3–5	3.2%	1.3%

* $p < 0.05$

In Exhibits 3.11 and 3.12, the analyses are repeated using race/ethnicity as the classification variable. The distribution in Exhibit 3.11 is statistically significant and shows that there were more Hispanics than blacks or whites who abstained from marijuana use in any of the index years. Interestingly, while use in one or two index years by Blacks and whites is slightly higher than for Hispanics, the largest differences are in the 3-5 year category, where Blacks and especially whites have much higher use rates than Hispanics. In Exhibit 3.12, there is no statistically significant racial/ethnic difference for the number of index years of cocaine use.

Exhibit 3.11: Number of Times Used Marijuana By Race/Ethnicity*

Number of Index Years	Hispanic	Black	White
Used Marijuana			
0	63.8%	57.3%	56.2%
1	18.3%	20.3%	19.3%
2	9.4%	10.4%	9.3%
3–5	8.4%	12.1%	15.2%

* $p < 0.05$

Exhibit 3.12: Number of Times Used Cocaine By Race/Ethnicity

Number of Index Years	Hispanic	Black	White
Used Cocaine			
0	82.3	83.2	79.9
1	11.3	10.9	12.5
2	4.5	4.1	5.1
3–5	2.0	1.7	2.6

COMPARISONS OF DRUG USE OVER TIME

Exhibits 3.13 through 3.18 present summary data about the number of years of drug use, and the additional years that drugs were used. Exhibits 3.13 and 3.14 present comparisons between drug use for adjacent study years. When marijuana use is considered over the entire study period, 53 percent of those who reported marijuana use in one time period did not use in the following study year. Further, 94 percent of those who did not use marijuana in the prior study period did not use in the following study year.

For cocaine, almost three-quarters (71 percent) of those who used in the prior time period did not use in the subsequent study year. In a similar pattern to marijuana, 97 percent of those who did not use cocaine in the prior time period did not use in the following study period. It seems that abstinence from drugs for one study year may be indicative of a substantial longer term effect.

Over the entire study period, the majority of respondents reported no drug use, as presented in Exhibit 3.15. Overall, 58.1 percent, 81.4 percent, and 96.6 percent of respondents reported no marijuana, cocaine, or crack use, respectively, in the five reporting periods. Only a small minority of individuals reported drug use in all five years (3.2 percent for marijuana, 0.2 percent for cocaine, and 0.2 percent for crack). Note that data on crack use were collected in 1992, 1994, and 1998 only.

Exhibit 3.13: Marijuana Use Between Adjacent Survey Periods

		First Time Period	
		Used Marijuana	Did Not Use Marijuana
Following Time Period	Used Marijuana	46.6%	6.2%
	Did Not Use Marijuana	53.4%	93.8%

Exhibit 3.14: Cocaine Use Between Adjacent Survey Periods

		First Time Period	
		Used Cocaine	Did Not Use Cocaine
Following Time Period	Used Cocaine	28.6%	3.0%
	Did Not Use Cocaine	71.4%	97.1%

Exhibit 3.15: Number of Years Used Drug

Number of Years Used	Marijuana		Cocaine		Crack*	
	Percentage	Count	Percentage	Count	Percentage	Count
0	58.1%	4,413	81.4%	6,185	96.6%	7,335
1	19.2%	1,459	11.7%	889	2.7%	203
2	9.6%	731	4.8%	363	0.6%	46
3	6.0%	453	1.3%	102	0.2%	13
4	3.9%	297	0.6%	45	NA	
5	3.2%	244	0.2%	13	NA	
Total	100%	7,597	100%	7,597	100%	7,597

*Data for use of crack are available only for 1992, 1994 and 1998.
 Note: Percentages may not sum to 100 due to rounding.

Exhibits 3.16 and 3.17 display the number of *additional* years that drug users continue their use of drugs, given drug use in one study year. Among those who reported marijuana use in any one year (Exhibit 3.16), almost half (46 percent) never again reported use in the other study years. However, 23 percent of individuals who used marijuana in one study year used in a second study year. Two-fifths (42 percent) of those who used marijuana in two study years did not report additional use.

Exhibit 3.17 presents the percentage of cocaine users who used cocaine additional years. Almost two-thirds (63 percent) of individuals who reported cocaine use in one study year did not report cocaine use in other study years. About a quarter (26 percent) of those who used cocaine in one

Exhibit 3.16: Percentage Who Used Marijuana Additional Years by Number of Years Used Marijuana

Additional Years Used	Number of Years Used			
	1	2	3	4
0	46%	42%	46%	55%
1	23%	26%	30%	45%
2	14%	17%	25%	NA
3	9%	14%	NA	NA
4	8%	NA	NA	NA
Total	100%	99%	101%	100%

Note: The number of years used is computed by adding each year of use for survey years 1984, 1988, 1992, 1994 and 1998.

Note: Percentages may not sum to 100 due to rounding.

Exhibit 3.17: Percentage Who Used Cocaine Additional Years by Number of Years Used Cocaine

Additional Years Used	Number of Years Used			
	1	2	3	4
0	63%	69%	64%	78%
1	26%	20%	28%	22%
2	7%	9%	8%	NA
3	3%	2%	NA	NA
4	1%	NA	NA	NA
Total	100%	100%	100%	100%

Note: The number of years used is computed by adding each year of use for survey years 1984, 1988, 1992, 1994, and 1998.

study year used in a second year. Two-thirds (69 percent) of those who used in two study years did not report additional use.

Exhibit 3.18 shows that almost all (96.6 percent) of respondents who did not use marijuana did not use cocaine, while more than half of those who used marijuana three or more study years also used cocaine.

COMPARISON BETWEEN NLSY AND NATIONAL HOUSEHOLD SURVEY ON DRUG ABUSE (NHSDA)

The discussion of drug use over time using a longitudinal survey leads to a different picture of drug use than is usually seen using a cross-sectional survey. Each survey type provides crucial information. While the cross-sectional survey provides an excellent snapshot of a specific time period, without the burden of interviewing and reinterviewing the same individuals over many years, longitudinal data permit inferences regarding change over time, as the information about each survey participant is sequenced along the years these individuals were interviewed.

This section of the report compares and contrasts two premiere data sets: the National Longitudinal Survey on Youth (the data used in this report) and the National Household Survey on Drug Abuse, sponsored by the Substance Abuse and Mental Health Services Administration of the Department of Health and Human Services. Exhibit 3.19 presents a comparison between these surveys. Except for 1985, we compare the relevant NLSY and NHSDA during the same year. When the NLSY was conducted in 1984, the NHSDA was not conducted that year. Instead, we used the 1985 NHSDA, the most proximate survey year, as the comparison with the 1984 NLSY. To compare the surveys, we used the same age cohorts for each.

The bulk of differences between the two nationally representative surveys were not statistically significant. The only differences were reports of higher rates of cocaine use (both lifetime and last year report) in the 1985 NHSDA and higher reports of lifetime marijuana use in the 1994 and 1998 NLSY than the NHSDA.

The NLSY data provide information on a cohort of 12,686 individuals and therefore allows us to investigate the following: (1) who uses drugs over extended periods (15 years); (2) who becomes a heavy drug user, and what variables predict such a phenomenon; (3) which individuals abstain or never use drugs, and what variables might help explain this phenomenon; and (4) which individuals use drugs intermittently, and why. The longitudinal data set also enables us to investigate the relationship between drug use in one period and a later period. In this way, longitudinal data provide advantages over cross-sectional data.

Exhibit 3.18: Number of Years Used Cocaine By Number of Years Used Marijuana*

Number of Years Used Cocaine	Number of Years Used Marijuana											
	0		1		2		3		4		5	
0	96.6%	4,261	77.5%	1,130	56.8%	415	44.4%	201	35.0%	104	30.3%	74
1	2.8%	125	17.6%	256	30.0%	219	30.7%	139	31.0%	92	23.8%	58
2	0.4%	19	4.2%	61	11.4%	83	17.7%	80	22.2%	66	22.1%	54
3	+		0.7%	10	+		5.7%	26	6.7%	20	12.3%	30
4	+		+		+		+		3.7%	11	8.6%	21
5	+		+		+		+		+		+	
Total	99.8%	4,413	100%	1,459	98.2%	731	98.5%	453	98.6%	297	97.1%	244

* Survey years 1984, 1988, 1992, 1994, 1998.

+ Fewer than 10 cases.

Note: Percentages may not sum to 100 due to rounding.

Exhibit 3.19: Comparison between NHSDA and NLSY for Lifetime and Past Year Marijuana and Cocaine Use 1984–1998

Age Group	NHSDA Year	NHSDA NATIONAL ESTIMATE (%)				NLSY Year	NLSY NATIONAL ESTIMATE (%)			
		Lifetime Marijuana	Past-Year Marijuana	Lifetime Cocaine	Past-Year Cocaine		Lifetime Marijuana	Past-Year Marijuana	Lifetime Cocaine	Past-Year Cocaine
20–27	1985	63.1	34.4	26.2	15.6	1984	61.0	31.8	15.2	10.0
23–30	1988	63.8	20.9	26.7	10.4	1988	62.9	21.0	26.8	9.7
27–34	1992	58.8	14.1	25.5	4.9	1992	55.1	12.2	23.8	3.6
29–36	1994	53.7	11.0	23.3	3.1	1994	61.4	13.4	23.8	3.4
33–40	1998	51.3	9.3	22.4	2.5	1998	56.3	9.2	21.7	2.1

NOTES: Due to changes in the questionnaire design in 1994, data from previous survey years are not directly comparable to 1994 estimates and subsequent survey years. The NLSY estimate is developed through analysis of the population described in this chapter (i.e., those individuals who responded to every one of the surveys in 1984, 1988, 1992, 1994, and almost always 1998).

All differences between NLSY and NHSDA are nonsignificant at the 0.05 level, except lifetime marijuana use in 1994 and 1998 (NLSY is significantly higher) and lifetime and past-year cocaine use in 1985 (NHSDA is significantly higher).

STRENGTHS OF LONGITUDINAL SURVEYS

Survey data can be collected to describe events over time through both prospective panel (longitudinal) surveys or retrospective cross-sectional means. Prospective data (e.g., NLYS79) are collected by taking repeated measures in a set of panel follow-ups. Retrospective data are generally collected in a single interview, using event history (calendar) methods to assist memory. The strengths of longitudinal surveys are their ability to measure transitions, changes, and the order of events as they occur. Longitudinal data collection provides more accurate data than cross-sectional data for long-term measures subject to response error caused by such things as recall and telescoping

A problem with retrospective cross-sectional data collection is the issue of recall accuracy.¹³ Also, given time limitations, the brief interview period of a single cross-sectional survey permits comparatively less detail on timing and change over time than is available from several interviews from a panel study. As a result, longitudinal panel data provides better answers than cross-sectional data to some questions, particularly because of lower recall error and better information on the order of events. For example:

- What is the relationship between drugs and crime and the movement from one to the other?
- Who becomes a heavy drug user? What variables predict heavy drug use?
- Who never uses drugs? Who uses drugs only a few times? Can we predict either?
- To what extent is drug use at one time period associated with use at another time? What explains the association?
- What is the relationship between exposure to prevention messages and drug use/criminal behavior?
- What are the transitions (initiation and cessation) of drug use and what is the duration of use?
- What is the relationship between unemployment and drug use? What explains the movement from one to the other? Which is prior?
- What is the effect of events (e.g., becoming unemployed or getting married or divorced) on initiation, cessation, or changes in drug use? Cross-sectional data collection from a single slice in time is less likely to learn about the point when events occur relative to changes in behavior.
- If important unobserved characteristics influence the outcome variable, the regression estimates will be biased. Longitudinal data allow us to obtain unbiased estimates through the use of fixed effects models that “net out” characteristics such as personality traits.

Summary

- Respondents included in the study share more in common with respondents excluded from the study. However, respondents included in the study report more drug use than those excluded from the study.

¹³ R. A. Johnson, D. R. Gerstein, K. A. Rasinski, “Adjusting Survey Estimates for Response Bias: An Application to Trends in Alcohol and Marijuana Use,” *Public Opinion Quarterly*, v. 62, no.3, 1998.

- A majority of individuals reported no drug use during the five study periods. Overall, 58.1 percent of interviewed individuals reported no marijuana use, 81.4 percent reported no cocaine use, and 96.6 percent reported no crack use (crack was only collected for the three study periods from 1992 to 1998).
- Very few individuals consistently reported drug use in all five years: for marijuana and cocaine, 3.2 percent and 0.2 percent reported use from 1984 to 1998, respectively. For crack, 0.2 percent reported use from 1992 to 1998.
- About half of marijuana users who used in one study period reported marijuana use in the subsequent study period. Only one of 16 who did not use marijuana in one study period used marijuana in the subsequent period.
- About one-quarter of cocaine users who used in one study period reported cocaine use in the subsequent study period. Only one of 32 who did not use cocaine in one study period used cocaine in the subsequent period.
- Of those who used in one study period, most did not use again. Overall, 19.2 percent, 11.7 percent, and 2.7 percent of individuals reported marijuana, cocaine or crack use, respectively, only in one year. Among those who reported marijuana use in any one year, almost half (46 percent) never reported use again. Among cocaine users, almost two-thirds (63 percent) of those who reported use in any one year never reported use again.
- More females than males reported no marijuana or cocaine use and more males than females reported marijuana or cocaine use during one or more of the index years. More Hispanics reported no marijuana use and more blacks reported marijuana use in one or two index years, whereas more whites reported marijuana use in three to five of the index years.
- In terms of poly-drug use, almost all of the interviewed individuals who did not report marijuana use did not report cocaine use (96.6 percent). More than half of those who reported marijuana use in three or more years also used cocaine.
- When compared to the similar age group in the same year of NHSDA data (except that there was no 1984 NHSDA), almost no significant differences were found in lifetime and last-year use of marijuana and cocaine. The 1985 NHSDA had significantly higher lifetime and past-year cocaine use, while the 1994 and 1998 NLSY showed significantly higher lifetime marijuana use.

4. HEAVY DRUG USE

Introduction

In this chapter, we analyze drug use, specifically heavy cocaine use. Our analysis is separated into the following components:

- Age of first use
- Prediction of heavy cocaine use
- Correlates of number of years of drug use
- Summary

Age of First Use

Our analysis also explored the relationships of heavy cocaine use and substance use by age. In particular, we examined the ages at which NLSY youth first smoked a cigarette, started smoking daily, started drinking twice a week, first smoked marijuana, first tried cocaine, and first tried crack cocaine. For a set of age ranges for which we have sufficient data, we report the percentage of heavy cocaine users among those who tried the drug by each particular age compared to the percentage of heavy cocaine users among those who had not tried the drug by that particular age. Our results are presented in Exhibits 4.1 through 4.6.

Exhibit 4.1 shows that 1,236 youth had smoked their first cigarette by age 10 (7,469 had not). Out of these 1,236 youth, 110 (8.90 percent) satisfied our conditions as a heavy cocaine user. This compares to a heavy cocaine user rate of only 5.73 percent among those who had not smoked their first cigarette by age 10. The highest rate of heavy cocaine users was among those who smoked their first cigarette at age 11 (43 out of 298 = 14.1 percent). After age 11, the longer a youth waited to smoke their first cigarette, the less likely they were to become a heavy cocaine user. This is shown by the declining percentages in the YES column as age increases. The declining percentages in the NO column also show that the longer a youth waits to smoke their first cigarette, the lower the chance that the youth will become a heavy cocaine user. For example, even if a youth has not started smoking by age 11, there is still a 5 percent chance that s/he will become a heavy cocaine user. This risk is less than 3 percent if the youth has not started smoking by age 17.

The ratio of percentages is also a measure of how much smoking cigarettes is associated with heavy cocaine use. Youth who have started smoking by age 10 are about 55 percent more likely to become heavy cocaine users as those who did not start smoking by age 10 ($8.90\% / 5.73\% = 1.55$). This ratio steadily increases, such that youth who have started smoking by age 17 are 2.5 times as likely (150 percent more likely) to become heavy cocaine users as those who have not started smoking by age 17. The increasing ratios are really just a facet of this particular presentation, which is designed to allow a comparison at any age cut-off. While the table shows that those who have started smoking by age 18 still have a 7.25 percent chance of becoming a heavy drug user, those who start at age 18 only have a 4.07 percent chance ($15/369$) of becoming a heavy drug user, which is closer to the risk of those who have not smoked by age 18 (2.78 percent) than it is to those who started by age 16 (7.62 percent).

Exhibit 4.1: Heavy Cocaine Use among those who had/had not smoked first cigarette by given age

	YES	NO
Used by age 10?	8.90% (110/1236)	5.73% (428/7469)
Used by age 11?	9.97% (153/1534)	5.37% (385/7171)
Used by age 12?	9.82% (241/2454)	4.75% (297/6251)
Used by age 13?	9.12% (301/3301)	4.39% (237/5404)
Used by age 14?	8.56% (356/4160)	4.00% (182/4545)
Used by age 15?	7.93% (397/5008)	3.81% (141/3697)
Used by age 16?	7.62% (444/5826)	3.27% (94/2879)
Used by age 17?	7.44% (465/6253)	2.98% (73/2452)
Used by age 18?	7.25% (480/6622)	2.78% (58/2083)
Ever Reported?	7.08% (494/6982)	2.55% (44/1723)

Exhibit 4.2 shows that 464 youth had started smoking daily by age 12 (8,188 had not). Out of these 464 youth, 62 (13.36 percent) satisfied our conditions as heavy cocaine users. This compares to a heavy cocaine user rate of only 4.96 percent among those who had not started smoking daily by age 12. The highest rate of heavy cocaine users was among those who started smoking daily by age 12. After age 12, the decreasing percentages in the YES column as age increases indicates that the longer a youth waits to start smoking daily, the less likely s/he is to become a heavy cocaine user (though it does seem to stop decreasing after age 18). The declining percentages in the NO column also show that the longer a youth waits to smoke daily, the lower the chance that the youth will become a heavy cocaine user. For example, even if a youth has not started smoking daily by age 12, there is still a 5 percent chance that s/he will become a heavy cocaine user. This risk is less than 3 percent if the youth has not started smoking daily by age 18.

The ratio of percentages is also a measure of the extent to which smoking cigarettes daily is associated with heavy cocaine use. Youth who start smoking daily by age 12 are about 2.7 times as likely to become heavy cocaine users as those who have not ($13.36\%/4.96\% = 2.69$). This ratio stays constant until it rises consistently after age 16, such that youth who have started smoking daily by age 21 are almost four times as likely ($8.85\%/2.27\% = 3.90$ times as likely) to become heavy cocaine users as those who have not started smoking daily by age 21. In this case, the increasing ratios are associated with a decrease in the NO column percentages, while the YES column percentages stay roughly constant (just under 9 percent). Those who start smoking daily at the age of 19, 20, or 21 still have a 8.15 percent (49/601) chance of becoming a heavy cocaine user, which is quite close to the percentage for those who started by age 18 (8.97 percent), and is much higher than those who have not started by age 21 (2.27 percent).

Exhibit 4.2: Heavy Cocaine Use among those who had/had not started smoking daily by given age

	YES	NO
Used by age 12?	13.36% (62/464)	4.96% (406/8188)
Used by age 13?	12.56% (100/796)	4.68% (368/7856)
Used by age 14?	11.01% (136/1235)	4.48% (332/7417)
Used by age 15?	10.58% (196/1853)	4.00% (272/6799)
Used by age 16?	9.81% (252/2570)	3.55% (216/6082)
Used by age 17?	9.48% (286/3018)	3.23% (182/5634)
Used by age 18?	8.97% (316/3522)	2.96% (152/5130)
Used by age 19?	8.78% (331/3768)	2.81% (137/4884)
Used by age 20?	8.83% (349/3953)	2.53% (119/4699)
Used by age 21?	8.85% (365/4123)	2.27% (103/4529)
Ever Reported?	8.56% (392/4582)	1.87% (76/4070)

Exhibit 4.3 shows that there are 170 youth in our sample who had started drinking alcohol twice a week by age 12 (8,618 had not). Out of these 170 youth, 30 (17.65 percent) satisfied our conditions as a heavy cocaine user. This compares to a heavy cocaine user rate of only 5.96 percent among those who had not started drinking alcohol twice a week by age 12. The highest rate of heavy cocaine users was among those who started drinking alcohol twice a week by age 12, since after age 12, the percentages in YES column decrease as age increases, indicating that the longer a youth waited to start drinking alcohol twice a week, the less likely they were to become a heavy cocaine user. The declining percentages in the NO column show that the longer a youth waits to start drinking alcohol twice a week, the lower the chance that the youth will become a heavy cocaine user. For example, even if a youth has not started drinking alcohol twice a week by age 12, there is still an almost 6 percent chance that s/he will become a heavy cocaine user. This risk is less than 2 percent if the youth has not started drinking alcohol twice a week by age 21.

The ratio of percentages is also a measure of the extent to which drinking alcohol twice a week leads to heavy cocaine use. Youth who have started drinking alcohol twice a week by age 12 are about 3.0 times as likely to become heavy cocaine users as those who have not started drinking alcohol twice a week by age 12 ($17.65\%/5.96\% = 2.96$). This ratio stays constant until it rises consistently after age 18, such that youth who have started drinking alcohol twice a week by age 21 are almost four times as likely ($7.32\%/1.95\% = 3.75$ times as likely) to become heavy cocaine users as those who have not started drinking alcohol twice a week by age 21. The increasing ratios are really just a facet of this particular presentation, which is designed to allow a comparison at any age cut-off. While the table shows that those who have started drinking alcohol twice a week by age 18 still have a 7.98 percent chance of becoming a heavy drug user, those who start at age 18 only have a 4.35 percent chance ($81/1863$) of becoming a heavy cocaine user, which is closer to the risk of those who have not started

drinking twice or more per week by age 18 (2.60 percent) than it is to those who started by age 16 (11.58 percent).

Exhibit 4.3: Heavy Cocaine Use among those who had/had not started drinking 2+/week by given age

	YES	NO
Used by age 12?	17.65% (30/170)	5.96% (514/8618)
Used by age 13?	16.76% (57/340)	5.76% (487/8448)
Used by age 14?	15.94% (99/621)	5.45% (445/8167)
Used by age 15?	14.96% (190/1270)	4.71% (354/7518)
Used by age 16?	11.58% (303/2617)	3.91% (241/6171)
Used by age 17?	9.67% (387/4003)	3.28% (157/4785)
Used by age 18?	7.98% (468/5866)	2.60% (76/2922)
Used by age 19?	7.65% (492/6434)	2.21% (52/2354)
Used by age 20?	7.46% (502/6728)	2.04% (42/2060)
Used by age 21?	7.32% (508/6942)	1.95% (36/1846)
Ever Reported?	7.24% (510/7045)	1.95% (34/1743)

Exhibit 4.4 shows that there are 358 youth had started smoking marijuana by age 12 (8,321 had not). Out of these 358 youth, 84 (23.46 percent) satisfied our conditions as a heavy cocaine user. This compares to a heavy cocaine user rate of only 5.34 percent among those who had not started smoking marijuana by age 12. The highest rate of heavy cocaine users was among those who started smoking marijuana by age 12, since after age 12, the decreasing percentages in YES column as age increases indicate that the longer a youth waited to start smoking marijuana, the less likely s/he was to become a heavy cocaine user. The declining percentages in the NO column also show that the longer a youth waits to start smoking marijuana, the lower the chance that the youth will become a heavy cocaine user. For example, even if a youth has not started smoking marijuana by age 12, there is still a 5 percent chance that s/he will become a heavy cocaine user. This risk is just over 1 percent if the youth has not started smoking marijuana by age 21.

The ratio of percentages is also a measure of the extent to which smoking marijuana is associated with heavy cocaine use. Youth who have started smoking marijuana by age 12 are about 4.4 times as likely to become heavy cocaine users as those who have not started smoking marijuana by age 12 ($23.46\% / 5.34\% = 4.39$). This ratio varies around 4.5 until it eventually rises consistently after age 17, such that youth who have started smoking marijuana by age 21 are almost eight times as likely ($8.11\% / 1.05\% = 7.72$ times) to become heavy cocaine users as those who have not started smoking marijuana by age 21. The increasing ratios are really just a facet of this particular presentation, which is designed to allow a comparison at any age cut-off. While the table shows that those who have started smoking marijuana by age 18 still have a 8.88 percent chance of becoming a heavy cocaine user, those who start at age 18 only have a 3.81 percent chance (34/892) of becoming a

heavy cocaine user, which is closer to the risk of those who have not smoked marijuana by age 18 (1.81 percent) than it is to those who started by age 16 (11.37 percent).

Exhibit 4.4: Heavy Cocaine Use among those who had/had not used marijuana by given age

	YES	NO
Used by age 12?	23.46% (84/358)	5.34% (444/8321)
Used by age 13?	19.74% (149/755)	4.78% (379/7924)
Used by age 14?	16.12% (211/1309)	4.30% (317/7370)
Used by age 15?	13.78% (295/2141)	3.56% (233/6538)
Used by age 16?	11.37% (385/3385)	2.70% (143/5294)
Used by age 17?	9.92% (432/4354)	2.22% (96/4325)
Used by age 18?	8.88% (466/5246)	1.81% (62/3433)
Used by age 19?	8.47% (481/5679)	1.57% (47/3000)
Used by age 20?	8.23% (492/5980)	1.33% (36/2699)
Used by age 21?	8.11% (502/6191)	1.05% (26/2488)
Ever Reported?	7.76% (513/6607)	0.72% (15/2072)

Exhibit 4.5 shows that 137 youth had started using cocaine by age 15 (8,392 had not). Out of these 137 youth, 48 (35.04 percent) satisfied our conditions as a heavy cocaine user. This compares to a heavy cocaine user rate of only 5.93 percent among those who had not started using cocaine by age 15. The highest rate of heavy cocaine users was among those who started using cocaine by age 15, since after age 15, the decreasing percentages in the YES column as age increases indicate that the longer a youth waited to start using cocaine, the less likely s/he was to become a heavy cocaine user. The declining percentages in the NO column also show that the longer a youth waits to start using cocaine, the lower the chance that the youth will become a heavy cocaine user. For example, even if a youth has not started using cocaine by age 15, there is still an almost 6 percent chance that s/he will become a heavy cocaine user. This risk is under 2 percent if the youth has not started using cocaine by age 24.

The ratio of percentages is also a measure of how much using cocaine is associated with heavy cocaine use. Youth who used cocaine by age 15 are almost six times as likely to become heavy cocaine users as those who have not started using cocaine by age 15 ($35.04\%/5.93\% = 5.91$). This ratio actually decreases to about 5 until it starts increasing after age 19, such that youth who have started using cocaine by age 25 are almost eleven times as likely ($15.61\%/1.46\% = 10.69$ times as likely) to become heavy cocaine users as those who have not started using cocaine by age 25. In this case, the increasing ratios are associated with a decrease in the NO column percentages, while the YES column percentages stay more constant (around 16 percent). Those who start using cocaine at the age of 22, 23, 24, or 25 still have a 10.93 percent (99/906) chance of becoming a heavy cocaine user, which is closer to the percentage for those who started by age 21 (17.66 percent) than those who have not started by age 21 (2.79 percent).

Exhibit 4.5: Heavy Cocaine Use among those who had/had not used cocaine by given age

	YES	NO
Used by age 15?	35.04% (48/137)	5.93% (498/8392)
Used by age 16?	25.57% (79/309)	5.68% (467/8220)
Used by age 17?	25.58%(132/516)	5.17% (414/8013)
Used by age 18?	21.72% (199/916)	4.56% (347/7613)
Used by age 19?	20.32% (252/1240)	4.03% (294/7289)
Used by age 20?	18.50% (314/1697)	3.40% (232/6832)
Used by age 21?	17.66% (366/2073)	2.79% (180/6456)
Used by age 22?	16.67% (393/2357)	2.48% (153/6172)
Used by age 23?	16.29% (416/2554)	2.18% (130/5975)
Used by age 24?	15.67% (431/2750)	1.99% (115/5779)
Used by age 25?	15.61% (465/2979)	1.46% (81/5550)
Ever Reported?	15.44% (521/3374)	0.48% (25/5155)

Exhibit 4.6 shows that there were 107 youth who had started using crack by age 18 (8,204 had not). Out of these 107 youth, 25 (23.36 percent) satisfied our conditions as a heavy cocaine user. This compares to a heavy cocaine user rate of only 5.28 percent among those who had not started using crack by age 18. Unlike the other variables, the percentages in the YES column do not monotonely decrease as age increases. The highest rate of heavy cocaine users was among those who started using crack between the ages of 18 and 21 ($46/127 = 36.22$ percent), but the percentage of crack users who become heavy cocaine users stays around 30 percent, even for those who start after age 30. The declining percentages in the NO column also show that the longer a person waits to start using crack, the lower the chance that the youth will become a heavy cocaine user. For example, even if a youth has not started using crack by age 18, there is still a 5 percent chance that s/he will become a heavy cocaine user. This risk is only about 3 percent if one does not start using crack by age 30.

The ratio of percentages is also a measure of how much smoking crack is associated with heavy cocaine use. Youth who have started using crack by age 18 are 4.4 times as likely to become heavy cocaine users as those who have not started using crack by age 18 ($23.36\%/5.28\% = 4.42$). This ratio steadily increases, such that youth who have started using crack by age 30 are 9.3 times as likely ($28.72\%/3.08\% = 9.32$) to become heavy cocaine users as those who have not started using crack by age 30. In this case, the increasing ratios are associated with a decrease in the NO column percentages, while the YES column percentages stay roughly constant (around 30 percent). Those who start using crack have a 30 percent chance of becoming heavy cocaine users no matter what age they start, while those who do not start by a certain age have their likelihood of becoming a heavy cocaine user decrease steadily with age.

Exhibit 4.6: Heavy Cocaine Use among those who had/had not used crack cocaine by given age

	YES	NO
Used by age 18?	23.36% (25/107)	5.28% (431/8204)
Used by age 21?	30.34% (71/234)	4.79% (385/8032)
Used by age 24?	30.19%(112/371)	4.36% (344/7895)
Used by age 27?	27.84% (169/607)	3.75% (287/7659)
Used by age 30?	28.72% (226/787)	3.08% (230/7479)
Ever Reported?	29.12% (272/934)	2.51% (184/7332)

It is not surprising that the largest differentials (rates of future heavy cocaine use for users are five or more times those for non-users) are among cocaine (Exhibit 4.5) and crack cocaine use (Exhibit 4.6), since these form the criteria for our heavy cocaine users. However, marijuana users are generally about four times or more as likely to become heavy cocaine users as those who have not tried marijuana at a young age (Exhibit 4.4). Those drinking alcohol twice a week by a certain age are generally about three times or more as likely to become heavy cocaine users as those who have not started drinking at least twice a week (Exhibit 4.3). Interestingly, smoking one cigarette by a certain age seems to at least double one's odds of becoming a heavy cocaine user (Exhibit 4.2), and smoking every day seems to triple the odds of one becoming a heavy cocaine user (Exhibit 4.1).

One key result in Exhibits 4.1–4.6 is that the longer youth wait to use any of these drugs, the less likely they are to become heavy cocaine users. This is shown by the “NO” column percentages decreasing uniformly through all six exhibits. However, this decrease is most significant for smoking the first cigarette, drinking alcohol, and smoking marijuana. For these three, later starters have rates of later heavy cocaine use closer to non-starters than to early starters. For smoking daily, using cocaine, or using crack, even late starters have significantly higher rates of later heavy cocaine use than non-starters.

Prediction of Heavy Cocaine Use

Regression is a useful tool to model data. One of our main goals is to find variables that can predict future heavy cocaine use. In order to stress the prediction, our emphasis for explanatory variables was on early round variables (generally in 1979 and 1980, once in 1984) so that we could test their predictive power on heavy cocaine use. Whether or not the youth is a heavy cocaine user is a binary variable, so we use logistic regression. Logistic regression is similar to ordinary linear regression, except that the dependent variable is dichotomous rather than continuous. For example, if y is 1 for a heavy cocaine user and 0 otherwise, then a logistic regression model posits the relationship between the explanatory variables and the probability that $y=1$:

$$\text{Log} \left[\frac{\text{Pr}(y = 1)}{\text{Pr}(y = 0)} \right] = \beta_0 + \sum \beta_i x_i$$

The probability itself cannot be used as the dependent variable because it is bounded by 0 and 1, whereas the right side of the equation is not. The log odds ratio is used because it is unbounded. The only difference from linear regression is the dependent variable. Any explanatory variables that were used for linear regression can be used for logistic regression.

In interpreting our logistic regression results, no claim is made for a direct causal relationship among the variables used in the proposed analysis. Rather, the study seeks to explain covariation among the variables to assess whether certain behaviors or characteristics tend to coincide with the presence of a particular outcome (i.e., heavy cocaine use) to a greater or lesser degree than do other variables or characteristics. Our goal is prediction.

Logistic regression generates “odds ratio” estimates for each predictor. Such estimates are readily interpretable probabilities that indicate how much more likely it is that an outcome would be observed if, all other elements being the same, the predictor occurs compared to when the predictor does not occur. For example, all other things being equal, an odds ratio would estimate how much more (or less) likely a man than a woman is to become a heavy cocaine user. An odds ratio above 1.0 means that the activity becomes more likely, while an odds ratio below 1.0 means that this activity becomes less likely.

Our modeling procedure followed a very programmatic path. We first selected 140 variables of possible interest from the NLSY79 data. All variables were treated as, or converted into categorical variables (where a goal was to have only a few categories). We examined the bivariate relationship between each of these variables and our heavy cocaine use dependent variable through chi-square tests of independence. These 140 items are presented in Appendix A sorted by item number (which are sorted by questionnaire year). Ninety-eight of these variables had a significant p value of less than 0.05, including 79 with a p value of less than 0.0001.

Clearly, these are too many independent variables to put into a logistic regression model, but we selected 26 variables for both their significance and their coverage of the important topics among the 140 variables. We ran several logistic regression stepwise procedures to determine the best model, using the SURvey DATA ANalysis (SUDAAN) software package. The best model shown below in Exhibit 4.7 included eight of these 26 variables.

The model in Exhibit 4.7 shows an R-square of only 0.08. However, logistic regression does not allow the full range of R-squares from zero to one. Instead, the R-square can be constrained quite strongly. Rescaling the R-square for logistic regression gives a fairer comparison to R-squares from linear regression. Cox and Snell's¹⁴ max-rescaled R-square is 0.24, which indicates a moderate explanatory model.

The odds ratios in Exhibit 4.7 are easily interpretable. The strongest factor is that those who smoked marijuana or hashish more than 50 times in the last year (in 1980) are 6.7 times as likely to be heavy cocaine users many years later than those who did not smoke any marijuana or hashish. This model shows that those who are smoking marijuana as youth are more likely to be heavy cocaine users as adults. This model also shows that those who smoked marijuana at least monthly (11–50 times) in 1980 were more than three times as likely to be heavy cocaine users than those who did not smoke

¹⁴ D. R. Cox and E. J. Snell, *Analysis of Binary Data*, 2nd Edition, London: Chapman and Hall, 1989.

any marijuana or hashish. Even occasional marijuana/hashish smokers (once or twice a year) were at least 50 percent more likely to become heavy cocaine users.

Exhibit 4.7: Correlates of Heavy Cocaine Use: Results of Logistic Regression Model

Observations used in analysis	8,033
R-square	0.08
Max-rescaled R-square	0.24
Independent Variables (Variables are significant at $p < 0.05$ level)	Odds Ratios
Smoked marijuana/hashish 51+ times in past year vs. none (1980)	6.67
Smoked marijuana/hashish 11–50 times in past year vs. none	3.23
Smoked marijuana/hashish 3–10 times in past year vs. none	2.38
Smoked marijuana/hashish 1–2 times in past year vs. none	1.54
Male vs. female (1979)	1.87
Had cigarette less than six months ago vs. never (1984)	1.41
Had cigarette more than six months ago vs. never	0.48
Has been suspended from school vs. has not (1980)	1.56
Significant amount of illegal income in last year vs. none (1980)	2.63
A little illegal income in last year vs. none	1.37
Attends religious services at least once a month vs. does not attend (1979)	0.67
Attends religious services infrequently vs. does not attend	0.76
Has sold hard drugs in past year vs. has not (1980)	2.00
Hispanic vs. non-black, non-Hispanic (1979)	1.59
Non-Hispanic black vs. non-black, non-Hispanic	1.61

Smoking cigarettes as young adults also has a strong relationship with heavy cocaine use. Current smokers in 1984 were 41 percent more likely to be heavy cocaine users than those who had never smoked. Interestingly, former smokers (those who had gone at least six months without smoking) were only half as likely to become heavy cocaine users as those who had never started smoking cigarettes.

Demographics also play a role. Males were 87 percent more likely than females to become heavy cocaine users, while non-Hispanic blacks and Hispanics were 61 percent and 59 percent, respectively, more likely than non-Hispanic, non-blacks to become heavy cocaine users.

Juvenile delinquent behavior is another valuable predictor of future heavy cocaine use. Youth who are suspended from school have a 56 percent greater chance in being a heavy cocaine user than those never suspended. Youth who received a substantial amount of illegal income were more than twice as likely to become heavy cocaine users. Those who reported “a little” illegal income were 37 percent more likely to become heavy cocaine users. Finally, those who sold hard drugs were twice as likely to become heavy cocaine users.

The final significant variable involves attendance at religious services. Those youth who attended religious services at least once per month in 1979 were 33 percent less likely to become heavy

cocaine users than those who did not attend religious services in 1979. Even those who attended religious services infrequently were 24 percent less likely to become heavy cocaine users than those who did not attend at all.

What is most striking about these relationships is that the explanatory variables were collected approximately five or more years (an average of 8 years) before the data were collected from which we classified the youth as heavy cocaine users. This suggests that it is possible to predict future heavy cocaine use several years in advance.

Correlates of Number of Years of Drug Use

We also ran linear regression analyses on the number of study years reported using marijuana, cocaine, and crack cocaine. The variables used in these regression analyses were not intended to be exhaustive of the data set or of other potentially meaningful relationships that could be assessed from this rich data resource. Instead, we used stepwise models with the same 26 variables we used in the heavy cocaine users logistic regression analysis to explore these three variables. Results for all three models are shown in Exhibit 4.8.

The modeling is much more effective for marijuana and cocaine than for crack. We believe this is due to the fact that there is much more marijuana use that can be modeled. While the weighted mean for marijuana is 1.15 years of use, the weighted mean for years of crack use is only 0.03, which is very small. None of the R-squares is as large as the max-rescaled R-square for heavy cocaine use (above), but marijuana is very close.

Some variables are important contributors to all three models. These are the same variables that are most important in predicting heavy cocaine use: smoking of marijuana/hashish, smoking status, and selling of hard drugs. Interestingly, ex-smokers (who are less likely to be heavy cocaine users) do tend to smoke more marijuana ($\beta = +0.15$), but not more cocaine ($\beta = 0.00$) or crack ($\beta = -0.01$).

Several variables were important for two of the three models. Those who had at least one episode of “binge drinking” (at least 5 alcoholic drinks at one time) in the last month were heavier users of marijuana and cocaine but not crack. Attending religious services is also significant and inversely related to number of years of marijuana and cocaine use, but not related to crack use.

Also significant is whether youth expect to be married within five years (possibly a proxy for whether the youth is in a committed relationship) and whether youth are satisfied with themselves; both are inversely related to number of years of drug use.

Race/ethnicity differences are not significant for cocaine, but African-Americans tend to use marijuana and crack for more years. However, Hispanics use marijuana for fewer years, and crack for about the same number of years as non-Hispanic whites. Youth who were referred to court-related counseling used cocaine and crack more years than those who were not referred to court-related counseling.

Several variables were significant in only one of the regressions. With regard to marijuana, males and those who had been suspended from school were users for more years, while for cocaine, those who expect to go to college are likely to use for slightly fewer years, but still use for more years than those who expect to get only a high school diploma (as expected, those who do not expect to finish

high school have the highest usage of cocaine). Finally, those who had at least one drink in the previous month and those with illegal income were likely to use cocaine for more years.

Exhibit 4.8: Correlates of Number of Years Using Marijuana, Cocaine, and Crack: Results of Ordinary Least Squares Regression Models

Observations used in analysis	Marijuana	Cocaine	Crack
R-square	4,490 0.23	4,391 0.18	8,293 0.04
Independent Variables (significant at $p < 0.05$ level)	values^a		
MEAN IN YEARS	1.15	0.40	0.03
Smoked marijuana/hashish more than 50 times in past year vs. none—(1980)	1.28	0.52	0.06
Smoked marijuana/hashish 11–50 times in past year vs. none	0.87	0.41	0.01
Smoked marijuana/hashish 3–10 times in past year vs. none	0.48	0.21	0.02
Smoked marijuana/hashish 1–2 times in past year vs. none	0.32	0.11	0.01
Had cigarette less than six months ago vs. never (1984)	0.53	0.14	0.03
Had cigarette more than six months ago vs. never	0.15	0.00	–0.01
Has sold hard drugs in past year vs. has not (1980)	0.39	0.26	0.05 ^b
Had at least five alcoholic drinks at one time in last month vs. did not (1983)	0.36	0.20	
Attends religious services at least once a month vs. does not attend (1979)	–0.14	–0.07	
Attends religious services infrequently vs. does not attend	–0.08	–0.04	
Hispanic vs. non-black, non-Hispanic (1979)	–0.06 ^b		0.01
non-Hispanic black vs. non-black, non-Hispanic	0.11 ^b		0.07
Has been referred to court-related counseling vs. has not (1980)		0.23	0.06
Expects to be married in next five years vs. expects to not be married (1979)	–0.30	–0.12	
Respondent is very satisfied with self vs. not satisfied (1980)	–0.11 ^b	–0.09	
Respondent is satisfied with self vs. not satisfied	–0.01 ^b	0.01	
Expects to go to college vs. expects to not finish high school (1979)		–0.02 ^b	
Expects highest grade to be HS diploma vs. expects to not finish HS		–0.10 ^b	
Had at least one alcoholic drink in the last month vs. had not (1980)		0.09	
Has been suspended from school vs. has not (1980)	0.13		
Male vs. female (1979)	0.19		
Significant amount of illegal income in last year vs. none (1980)		0.29	
A little illegal income in last year vs. none		0.12	

^a This effect is the additional years of drug use attributable to this level of the variable. Blank cells indicate variable was not significant in that regression.

^b Only significant at $p < 0.10$.

Summary

- Youth who wait longer before their first use of alcohol, cigarettes, marijuana, cocaine, or crack are less likely to become heavy cocaine users.
- Youth who first used cigarettes, alcohol, and marijuana late more closely resemble non-users with regard to their heavy cocaine use.
- However, those youth who started smoking cigarettes daily, using cocaine, or using crack late have heavy cocaine use percentages more like early users than non-users.
- Those who smoked marijuana more than 50 times as adolescents are more than six times as likely to become heavy cocaine users as those who did not smoke marijuana as adolescents. The odds ratio decreases as marijuana use in adolescence decreases.
- Young male drug users are almost twice as likely as female drug users to become heavy cocaine users.
- Those who were suspended from school are one and one-half times more likely to become heavy cocaine users as those who were not suspended from school.
- Individuals reporting a significant amount of illegal income as adolescents are more than two times more likely to become heavy cocaine users than those who had no illegal income as adolescents.
- Those who attended religious services at least twice a month are one third less likely to become heavy cocaine users as those who did not attend religious services. Attending infrequently is associated with a one-quarter lower likelihood of becoming a heavy cocaine user.
- Those selling hard drugs during adolescence are twice as likely to become heavy cocaine users as those who did not sell drugs as adolescents.

5. LONGITUDINAL SCREENING

Introduction

The final logistic regression model for heavy drug use above can be used as a screening tool to predict future heavy cocaine use. The variables in this final logistic regression model point out questions that could be used in future surveys or general population applications. The answers can then be input into this model to predict the likelihood of each respondent becoming a heavy drug user. It is also possible to use this likelihood to oversample the “most” likely future heavy drug users. As long as this is done on a probability basis, we can use weights to get a nationally representative and unbiased sample no matter how strong the oversampling is. Of course, the more oversampling there is, the more variability we would have in our nationally representative estimates using the weights (precision of the estimates is mathematically related to the variation in the weights). More oversampling, of course, also should lead to a larger sample of heavy cocaine users.

Screening the Sample

One goal was to develop a logistic regression model to help screen for a longitudinal survey that includes appropriate numbers of likely heavy drug users. As above, the screener questions would be designed with the model in mind, with the responses input to predict the likelihood of becoming a heavy drug user. The sample could be selected with a probability proportional to this estimated likelihood of becoming a heavy drug user. For example, since the odds ratio for males indicates a (roughly) doubled risk of becoming heavy drug users versus females, then males will be selected at twice the rate of females.

It is important to realize that age is a critical factor in any longitudinal survey that would study heavy drug use. The age of the youth need to be determined to balance the benefits and costs of starting with younger children who could develop into heavy drug users versus older children (e.g., teenagers and college-age) who could already be using heavy drugs. According to our analyses of age of first use, 14.2 percent of the NLSY sample had smoked a cigarette by age 10, and 10.7 percent had used cocaine by age 18.

While it will be difficult to ask about marijuana and other sensitive topics during a screening interview, there are several possible approaches. One straightforward approach is to have a screening interview only to determine age-eligible children and performing a short interview to ask the more sensitive questions as part of a further screener.

This would involve a lot of screening, so another suggestion is to try to “piggy-back” on another large screening sample. In this way, screening money could be saved if permission could be granted to simply target age-eligible children found by another project’s large screening operation. The 1997 Profiles of American Youth (PAY97) sponsored by the Department of Defense did just this in partnership with the 1997 National Longitudinal Survey of Youth (NLSY97) sponsored by the Bureau of Labor Statistics. NLSY97 (the newer cohort to NLSY79) screened a national sample of 90,000 housing units for 12- to 16-year-old youth. At the same time, PAY97 screened for 18- to 23-year-old youth and 10th- to 12th-grade students.

Selecting the Sample

Let us assume that screener data are available for all the variables in Exhibit 4.7: Number of times smoked marijuana/hashish in the past year, gender, race/ethnicity, smoking status (most recent cigarette), any school suspension history, amount of illegal income in last year, selling of hard drugs, and religious attendance. Taking the idea expressed above, a sample proportional to their predicted probability of becoming a heavy cocaine user can be selected.

In order to simplify this discussion, we use unweighted NLSY79 data. The rate of future heavy drug users among the NLSY79 respondents was $551/8,033 = 6.86$ percent. If we used an equal probability sample of 8,033 screener respondents (including the variables in Exhibit 4.7), we would then expect to have 551 future heavy cocaine users in our sample. However, by selecting our sample based on the screener variables shown to be related to future heavy cocaine use, we can increase our yield in a sample of 8,033. For example, the odds ratio for gender is 1.87, which suggests that the rate of future heavy cocaine users is 4.78 percent for females and $1.87 * 4.78 = 8.94$ percent for males (please note that the average of 4.78 percent and 8.94 percent is 6.86 percent). If instead of a 50/50 sample of males and females, we selected 87 percent more males (34.84 percent female), we would obtain 2,799 females (134 future heavy cocaine users) and 5,234 males (468 future heavy cocaine users). This would result in $468 + 134 = 602$ future heavy cocaine users (7.49 percent), an increase of over 9 percent over a gender-balanced sample. This assumes, of course, that the screening data contains at least 5,234 males.

A key idea is that the number of future heavy cocaine users can be increased further by increasing the oversampling of males. However, using only gender, we could not expect to exceed a percentage of future heavy cocaine users of 8.94 percent (which would consist of a 100 percent male sample).

Similar to gender, we could repeat the above analysis for all other screener variables. The most productive of these would surely be the number of times a respondent smoked marijuana or hashish in the last year, as shown by the odds ratio of 6.67 in Exhibit 4.7 for those who smoked marijuana or hashish at least 50 times in the last year. Examining this variable shows that 1,815 reported more than 50 uses in the last year, 817 reported 11–50 uses, 1,359 reported 3–10 uses, and 6,229 reported no use. The parameters in Exhibit 4.7 suggest that the rates of future heavy cocaine use are 20.22 percent (more than 50 uses), 9.79 percent (11–50 uses), 7.21 percent (3–10 uses), 4.67 percent (1–2 uses), and 3.03 percent (no use). Selecting the sample proportional to these predicted probabilities would result in a sample of (total sample of 8,033) 3,667 sample members with more than 50 uses (742 future heavy cocaine users), 799 with 11–50 uses (78 future heavy cocaine users), 980 with 3–10 uses (71 future heavy cocaine users), 701 with 1–2 uses (33 future heavy cocaine users), and 1,887 with 0 uses (57 future heavy cocaine users). This sample results in 980 future heavy cocaine users (an increase of 78 percent)—using only the one variable.

Using selection probabilities based on using all of the variables in Exhibit 4.7 will obviously lead to the possibility of even more future heavy cocaine users in the sample. Calculations, however, are highly dependent on the distribution of the screening sample across all cells for these variables, and are therefore not shown here. However, the above example for marijuana use shows that the rate of future heavy cocaine users can easily be increased from the NLSY97 rate of 6.86 percent to more than 10 percent using only this one variable in the selection of the sample.

Weighting the Sample

It should be noted that the survey need not be restricted to “likely” heavy drug users; “likely” heavy drug users would simply be overrepresented. For example, the above examples still include females and non-marijuana smokers in the sample. Therefore, this sample would still be nationally representative with the proper weights. Weights are commonly used to adjust for differential probabilities of selection. In fact, a Horvitz-Thompson¹⁵ estimator is simply a weighted mean where the weight is the reciprocal of the selection probability:

$$\hat{y} = \frac{1}{N} \sum \frac{y_i}{p_i},$$

where N is the population size, y_i is the observation, and p_i is the selection probability. Horvitz-Thompson estimators have been studied extensively, but their main benefit is that they are unbiased. In this case, this unbiased property results in nationally representative estimates.

Taking our example samples selected above, if males were selected with a probability (p_i) 87 percent greater than females, the base weight ($1/p_i$) for females would be 87 percent greater. Taking a very simple example, let’s assume that the sample of 2,799 females and 5,234 males was taken from a population of 25,000 females and 25,000 males. The selection probabilities in this case would be $2,799/25,000 = 11.20$ percent for the females and $5,234/25,000 = 20.94$ percent for the males. The base weight would be $1/0.1120 = 8.93$ for the females and $1/0.2094 = 4.78$ for the males. Since we would then expect 134 female and 468 male future heavy drug users, the above formula would imply our estimate of the future heavy drug user percentage (where 1= future heavy drug user, 0 = not) in the entire population of 50,000 males and females is:

$$\hat{y} = \frac{1}{50,000} \left(\frac{134 * 1}{.1120} + \frac{468 * 1}{.2094} \right) = .0686$$

This estimate of 6.86 percent matches our NLSY97 (unweighted) data, and implies 3,431 future heavy drug users in the theoretical population of 50,000 youth.

Conclusion

A longitudinal study would fill a currently large gap in our knowledge base of drug use. Currently, most data on drug use comes from cross-sectional surveys, from which it is very difficult to learn the temporal order of the factors that lead to drug use and abuse.

¹⁵ W. G. Cochran, *Sampling Techniques*, 3rd edition, New York: Wiley, 259-261, 1977.

APPENDIX A

Selected Variables of Interest from NLSY79 Data*

Item #	Question	Year	CHI^2	Df	P value	N (12686)
30	LIVED IN PRESENT RESIDENCE SINCE BIRTH? 79 INT	79	0.08	1	0.77	12685
65	HIGHEST GRADE COMPLETED BY R'S MOTHER 79 INT	79	12.47	6	0.05	12656
68	DID MOTHER/STEPMOTHER WORK FOR PAY? 1978	79	3.77	2	0.15	12030
70	DID MOTHER/STEPMOTHER WORK > 35 HOURS PER WEEK IN 1978?	79	1.94	1	0.16	7147
79	HIGHEST GRADE COMPLETED BY R'S FATHER 79 INT	79	16.89	6	0.01	12453
84	DID FATHER/STEPFATHER WORK > 35 HOURS PER WEEK IN 1978?	79	0.02	1	0.88	9124
91	NUMBER OF SIBLINGS 79 INT	79	5.34	6	0.50	12668
95	HIGHEST GRADE COMPLETED BY OLDEST SIBLING 79 INT	79	8.91	6	0.18	9694
96	1ST OR ONLY RACIAL/ETHNIC ORIGIN 79 INT	79	27.27	14	0.02	12524
105	FREQUENCY OF R'S RELIGIOUS ATTENDANCE 79 INT	79	61.58	5	0.00	12669
153	TOTAL NUMBER OF CHILDREN R EXPECTS TO HAVE 79 INT	79	18.88	5	0.00	12460
159	CURRENT SCHOOL ATTITUDE—FRIENDS CAN EASILY BE MADE	79	2.86	3	0.41	5730
160	CURRENT SCHOOL ATTITUDE—TEACHERS WILLING TO HELP	79	0.81	3	0.85	5726
161	CURRENT SCHOOL ATTITUDE—CLASSES ARE BORING 79 INT	79	2.56	3	0.46	5729
162	CURRENT SCHOOL ATTITUDE—R DOES NOT FEEL SAFE 79 INT	79	5.37	3	0.15	5730
163	CURRENT SCHOOL ATTITUDE—TEACHERS KNOW THEIR SUBJECTS	79	2.06	3	0.56	5726
164	CURRENT SCHOOL ATTITUDE—STUDENTS CAN "GET AWAY WITH"	79	19.47	3	0.00	5725
165	CURRENT SCHOOL ATTITUDE—SCHOOL WORK REQUIRES THOUGHT	79	13.20	3	0.00	5729
166	CURRENT SCHOOL ATTITUDE—FREEDOM TO PURSUE INTERESTS	79	3.51	3	0.32	5726
167	CURRENT SCHOOL ATTITUDE—SCHOOL OFFERS GOOD JOB	79	9.81	3	0.02	5686
168	ATTITUDE TOWARD CURRENT SCHOOL 79 INT	79	9.25	3	0.03	5709
171	REASON R LEFT SCHOOL (NOT ENROLLED) 79 INT	79	68.03	11	0.00	5231
172	HIGHEST GRADE ATTENDED 79 INT	79	46.98	5	0.00	12675
173	HIGHEST GRADE COMPLETED 79 INT	79	40.48	5	0.00	12676
175	IS CURRENT OR LAST SCHOOL GRADES 1–12 PUBLIC	79	1.11	1	0.29	12665
186	INT CHECK 79—IS R CURRENTLY ENROLLED IN GRADES 9–12?	79	0.09	1	0.76	12644
196	TYPE OF CURRENT OR LAST SCHOOL CURRICULUM IN GRADES 9–12	79	12.94	3	0.00	11526
234	HIGHEST GRADE COMPLETED R WOULD LIKE 79 INT	79	10.32	4	0.04	12624
235	HIGHEST GRADE COMPLETED R EXPECTS 79 INT	79	32.30	4	0.00	12565
236	HIGHEST GRADE COMPLETED R'S CLOSEST FRIEND WOULD LIKE 79	79	22.38	4	0.00	12366
270	PROBLEMS IN GETTING A GOOD JOB—TRANSPORTATION? 79 INT	79	2.56	1	0.11	10065
271	PROBLEMS IN GETTING A GOOD JOB—RACE DISCRIMINATION?	79	1.50	1	0.22	10058
272	PROBLEMS IN GETTING A GOOD JOB—NATIONALITY DISCRIMINATION?	79	0.11	1	0.74	10054
273	PROBLEMS IN GETTING A GOOD JOB—SEX DISCRIMINATION? 79	79	0.02	1	0.89	10057
274	PROBLEMS IN GETTING A GOOD JOB—AGE DISCRIMINATION? 79	79	15.05	1	0.00	10063
275	PROBLEMS IN GETTING A GOOD JOB—ENGLISH IS A FOREIGN LANG	79	0.38	1	0.54	10059
434	ACTIVITY MOST OF SURVEY WEEK CPS ITEM 79 INT	79	43.46	6	0.00	11465
507	R'S OPINION OF DEGREE OF DIFFICULTY IN FINDING A NEW JOB	79	0.43	2	0.81	6301
1451	INT CHECK 79—DOES R HAVE HEALTH LIMITATIONS—ITEMS 2A	79	0.00	1	0.94	12465
1709	WORK COMMITMENT—SEEK MORE EDUC IF UNABLE TO SUPPORT FAMILY	79	5.01	1	0.03	12560
1710	WORK COMMITMENT—GO ON WELFARE IF UNABLE TO SUPPORT FAMILY	79	2.50	1	0.11	12423
1711	WORK COMMITMENT—ENTER TRAINING IF UNABLE TO SUPPORT FAMILY	79	0.14	1	0.71	12571
1712	WORK COMMITMENT—APPLY 4 FOOD STAMPS IF UNABLE TO SUP FAM	79	3.99	1	0.05	12432
1713	WORK COMMITMENT—SHOPLIFT IF UNABLE TO SUPPORT FAMILY	79	31.97	1	0.00	12574
1714	WORK COMMITMENT—WOULD WORK IF COULD LIVE COMFORTABLY	79	1.15	1	0.28	12646
1716	MARITAL STATUS EXPECTATIONS—MARRIED OR NOT IN 5 YEARS	79	4.37	1	0.04	10453
1718	EDUCATIONAL EXPECTATIONS—IN SCHOOL OR NOT IN 5 YEARS	79	0.05	1	0.81	12237
1719	EMPLOYMENT EXPECTATIONS—WORKING OR NOT IN 5 YEARS 79	79	3.47	1	0.06	12413
1720	OCCUPATION EXPECTATIONS IN 5 YEARS (CENSUS 3 DIGIT) 79	79	28.99	13	0.01	12686

Item #	Question	Year	CHI^2	Df	P value	N (12686)
2147	R'S RACIAL/ETHNIC COHORT FROM SCREENER 79 INT	79	30.51	2	0.00	12686
2148	SEX OF R 79 INT	79	100.19	1	0.00	12686
2165	AGE OF R AT INTERVIEW DATE 79 INT	79	14.67	8	0.07	12686
2167	HIGHEST GRADE COMPLETED AS OF MAY 1 SURVEY YEAR	79	44.31	6	0.00	12680
2314	EVER BEEN SUSPENDED FROM SCHOOL? 80 INT	80	276.20	1	0.00	12140
2315	NUMBER OF TIMES R HAS BEEN SUSPENDED FROM SCHOOL 80 INT	80	351.73	6	0.00	12106
3035	SELF-ESTEEM 80 INT—I AM A PERSON OF WORTH 80 INT	80	3.89	2	0.14	12120
3036	SELF-ESTEEM 80 INT—I HAVE A NUMBER OF GOOD QUALITIES	80	1.02	2	0.60	12117
3037	SELF-ESTEEM 80 INT—I FEEL THAT I AM A FAILURE	80	5.76	2	0.06	12108
3038	SELF-ESTEEM 80 INT—I AM AS CAPABLE AS OTHERS 80 INT	80	1.60	2	0.45	12116
3039	SELF-ESTEEM 80 INT—I DO NOT HAVE MUCH TO BE PROUD	80	5.71	2	0.06	12104
3040	SELF-ESTEEM 80 INT—I HAVE A POSITIVE ATTITUDE 80 INT	80	0.04	2	0.98	12083
3041	SELF-ESTEEM 80 INT—I AM SATISFIED WITH MYSELF 80 INT	80	13.84	2	0.00	12103
3042	SELF-ESTEEM 80 INT—I WISH I HAD MORE SELF-RESPECT 80	80	2.46	3	0.48	12091
3043	SELF-ESTEEM 80 INT—I FEEL USELESS AT TIMES 80 INT	80	1.85	3	0.60	12114
3044	SELF-ESTEEM 80 INT—I SOMETIMES THINK I AM "NO GOOD"	80	3.98	3	0.26	12113
3046	ILLEGAL ACTIVITY 80 INT—TIMES RUN AWAY FROM HOME IN PAST YEAR	80	58.58	4	0.00	12141
3047	ILLEGAL ACTIVITY 80 INT—TIMES SKIPPED SCHOOL DAY IN PAST YEAR	80	100.08	6	0.00	3896
3048	ILLEGAL ACTIVITY —TIMES DRANK ALCOHOLIC BEVERAGES PAST YR	80	103.26	6	0.00	3886
3049	ILLEGAL ACTIVITY 80 INT—TIMES INTENTIONALLY DAMAGED IN PAST YR	80	150.90	4	0.00	11734
3050	ILLEGAL ACTIVITY 80 INT—TIMES FOUGHT AT SCHOOL OR WORK PAST YR	80	155.51	4	0.00	11800
3051	ILLEGAL ACTIVITY 80 INT—TIMES SHOPLIFTED IN PAST YEAR	80	217.08	4	0.00	11788
3052	ILLEGAL ACTIVITY 80 INT—TIMES STOLEN OTHER'S < \$50 IN PAST YR	80	114.03	4	0.00	11788
3053	ILLEGAL ACTIVITY 80 INT—TIMES STOLEN OTHER'S > \$50 IN PAST YR	80	195.29	3	0.00	11776
3054	ILLEGAL ACTIVITY —TIMES USED FORCE TO OBTAIN THINGS IN PAST YR	80	37.60	1	0.00	11794
3055	ILLEGAL ACTIVITY —TIMES SERIOUSLY THREATENED TO HIT/HIT PAST YR	80	97.55	4	0.00	11785
3056	ILLEGAL ACTIVITY 80 INT—TIMES ATTACKED W/INTENT TO HARM PAST YR	80	94.50	1	0.00	11792
3057	ILLEGAL ACTIVITY 80 INT—TIMES SMOKED MARIJUANA/HASHISH PAST YR	80	451.98	6	0.00	11722
3058	ILLEGAL ACTIVITY —TIMES USED OTHER DRUGS/CHEMICALS PAST YR	80	452.39	6	0.00	11698
3059	ILLEGAL ACTIVITY 80 INT—TIMES SOLD MARIJUANA/HASHISH PAST YR	80	408.35	6	0.00	11693
3060	ILLEGAL ACTIVITY 80 INT—TIMES SOLD HARD DRUGS IN PAST YR	80	182.66	1	0.00	11717
3061	ILLEGAL ACTIVITY 80 INT—TIMES ATTEMPTED TO "CON" PAST YR	80	99.96	4	0.00	11721
3062	ILLEGAL ACTIVITY 80 INT—TIMES TAKEN AUTO W/OUT OWNER PAST YR	80	60.60	1	0.00	11752
3063	ILLEGAL ACTIVITY 80 INT—TIMES BROKEN INTO A BUILDING PAST YR	80	116.16	1	0.00	11748
3064	ILLEGAL ACTIVITY 80 INT—TIMES KNOWINGLY SOLD/HELD PAST YR	80	168.04	1	0.00	11749
3065	ILLEGAL ACTIVITY 80 INT—TIMES AIDED IN GAMBLING PAST YR	80	35.25	1	0.00	11737
3066	ILLEGAL ACTIVITY —AMT OF TOTAL ILLEGAL INCOME IN PAST YEAR	80	315.99	2	0.00	11447
3067	EVER "STOPPED" BY POLICE FOR OTHER THAN MINOR TRAFFIC	80	139.20	1	0.00	12129
3068	TIMES "STOPPED" BY POLICE 80 INT (EXC MINOR TRAFFIC)	80	194.83	6	0.00	12103
3069	TIMES "STOPPED" BY POLICE IN PAST YEAR 80 INT (EXC MINOR TRAFFIC)	80	147.02	3	0.00	12111
3070	AGE OF R 1ST TIME "STOPPED" BY POLICE 80 INT	80	181.96	6	0.00	12111
3071	EVER CHARGED WITH ILLEGAL ACTIVITY? 80 INT (EXC MINOR TRAFFIC)	80	209.95	1	0.00	12136
3072	TIMES CHARGED WITH ILLEGAL ACTIVITY 80 INT (EXC MINOR TRAFFIC)	80	233.89	3	0.00	12132
3076	AGE AT TIME OF 1ST ILLEGAL ACTIVITY CHARGE 80 INT (EXC MINOR TRAF)	80	226.92	7	0.00	12129
3077	EVER CHARGED WITH ANY ILLEGAL ACTIVITY IN ADULT COURT?	80	0.88	1	0.35	1322
3078	EVER CONVICTED ON CHARGES OTHER THAN MINOR TRAFFIC	80	143.08	1	0.00	12130
3101	EVER BEEN REFERRED TO COURT-RELATED COUNSELING? 80 INT	80	167.70	1	0.00	12121
3105	INT CHECK 80—R EVER BEEN STOPPED, BOOKED/CHARGED,	80	208.59	1	0.00	12124
3109	EVER BEEN SENTENCED TO ANY TYPE OF CORRECTIONAL FACILITY	80	6.90	1	0.01	2749
7801	ALCOHOL USE—EVER HAD A DRINK? 82 INT	82	16.71	1	0.00	12122
7802	ALCOHOL USE—AGE WHEN STARTED DRINKING 82 INT	82	102.41	6	0.00	11111
7803	ALCOHOL USE—HAD ANY ALCOHOLIC BEVERAGES IN LAST MONTH?	82	49.62	1	0.00	12121
7804	ALCOHOL USE—FREQUENCY OF 6 OR MORE DRINKS IN LAST MONTH	82	159.22	6	0.00	7885
7805	ALCOHOL USE—FREQUENCY OF GOING TO BARS IN LAST MONTH	82	100.78	5	0.00	7887
7806	ALCOHOL USE—# OF DAYS DRANK ALCOHOL IN LAST WEEK 82	82	120.90	7	0.00	7882

Item #	Question	Year	CHI^2	Df	P value	N (12686)
7807	ALCOHOL USE—# OF CANS/BOTTLES OF BEER CONSUMED IN LAST WEEK	82	82.31	3	0.00	6147
7808	ALCOHOL USE—# OF GLASSES OF WINE CONSUMED IN LAST WEEK	82	5.05	2	0.08	6153
7809	ALCOHOL USE—# OF DRINKS WITH LIQUOR CONSUMED IN LAST WEEK	82	53.41	2	0.00	6154
7810	ALCOHOL USE—DRINKING EVER INTERFERED WITH SCHOOL	82	12.95	1	0.00	11087
7811	ALCOHOL USE—DRINKING EVER INTERFERED WITH JOB? 82	82	49.43	1	0.00	11071
7812	ALCOHOL USE—INT CHECK 83—ANYONE ELSE PRESENT WHEN	82	0.46	1	0.50	11042
10216	ALCOHOL USE—EVER HAD A DRINK? 83 INT	83	10.08	1	0.00	12209
10217	ALCOHOL USE—AGE WHEN STARTED DRINKING 83 INT	83	236.59	9	0.00	11407
10218	ALCOHOL USE—AGE WHEN 1ST STARTED DRINKING AT LEAST WEEKLY	83	176.90	11	0.00	9613
10219	ALCOHOL USE—HAD ANY ALCOHOLIC BEVERAGES IN LAST MONTH?	83	67.60	1	0.00	11417
10220	ALCOHOL USE—FREQUENCY OF 6 OR MORE DRINKS LAST MONTH	83	176.56	6	0.00	8055
10221	ALCOHOL USE—# OF DAYS DRANK ALCOHOL IN LAST MONTH 83	83	294.96	12	0.00	12206
10222	ALCOHOL USE—# OF DAYS HAD 1 DRINK IN LAST MONTH 83 INT	83	124.04	7	0.00	12203
10223	ALCOHOL USE—# OF DAYS HAD 2 DRINKS IN LAST MONTH 83 INT	83	112.54	7	0.00	12203
10224	ALCOHOL USE—# OF DAYS HAD 3 DRINKS IN LAST MONTH 83 INT	83	134.12	7	0.00	12203
10225	ALCOHOL USE—# OF DAYS HAD 4 DRINKS IN LAST MONTH 83 INT	83	225.57	7	0.00	12203
10226	ALCOHOL USE—# OF DAYS HAD 5 DRINKS IN LAST MONTH 83 INT	83	92.54	2	0.00	12203
10227	ALCOHOL USE—# OF DAYS HAD 6 OR MORE DRINKS IN LAST MONTH	83	227.13	8	0.00	12203
10229	ALCOHOL USE—FREQUENCY OF GOING TO BARS IN LAST MONTH 83	83	75.13	5	0.00	8054
10230	ALCOHOL USE—# OF DAYS HAD HANGOVER IN LAST MONTH 83 INT	83	162.94	4	0.00	12210
10231	ALCOHOL USE—# OF DAYS DRANK ALCOHOL IN LAST WEEK 83 INT	83	182.61	7	0.00	8061
10232	ALCOHOL USE—# OF CANS/BOTTLES OF BEER CONSUMED IN LAST WEEK	83	257.65	7	0.00	12204
10233	ALCOHOL USE—# OF GLASSES OF WINE CONSUMED IN LAST WEEK	83	148.82	3	0.00	12202
10234	ALCOHOL USE—# OF DRINKS WITH LIQUOR CONSUMED IN LAST WEEK	83	231.79	4	0.00	12202
10235	ALCOHOL USE—DRINKING EVER INTERFERED WITH SCHOOL WORK?	83	16.26	1	0.00	11343
10236	ALCOHOL USE—DRINKING EVER INTERFERED WITH JOB? 83 INT	83	45.90	1	0.00	11343
10237	ALCOHOL USE—INT CHECK 83—ANYONE ELSE PRESENT DURING SEC 12	83	3.87	2	0.14	12197
13948	DRUG USE—AGE WHEN 1ST TRIED A CIGARETTE? 84 INT	84	89.39	13	0.00	12030
13949	DRUG USE—MOST RECENT TIME HAD A CIGARETTE? 84 INT	84	122.73	3	0.00	12045
13950	DRUG USE—# OF CIGARETTES SMOKED PER DAY IN PAST 30 DAYS	84	8.62	5	0.13	5161
39142	DRUG USE—R SMOKED AT LEAST 100 CIGARETTES IN ENTIRE LIFE?	92	149.54	1	0.00	8575
39147	DRUG USE—# CIGARETTES SMOKED PER DAY (R SMOKES DAILY)	92	138.61	4	0.00	8929
50524	DRUG USE—R SMOKED AT LEAST 100 CIGARETTES IN ENTIRE LIFE?	94	184.19	1	0.00	8826
50530	DRUG USE—# CIGARETTES SMOKED PER DAY (R SMOKES DAILY)	94	191.55	6	0.00	8869
64297	DRUG USE—R SMOKED AT LEAST 100 CIGARETTES IN ENTIRE LIFE?	98	155.43	1	0.00	8333
64303	DRUG USE—# CIGARETTES SMOKED PER DAY (R SMOKES DAILY)	98	179.43	7	0.00	8387

*Variables in bold were used in regression modeling.



**EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF NATIONAL DRUG CONTROL POLICY
WASHINGTON, DC 20503**