# THE MONITORING THE FUTURE EIGHTH GRADE PANEL SURVEY DATA: SAMPLE DESIGN, ADJUSTMENTS FOR PANEL ATTRITION BIASES, AND ASSESSMENT OF MEASUREMENT BIAS

*Monitoring the Future Occasional Paper 55*

Jerald G. Bachman
Patrick M. O'Malley
Lloyd D. Johnston
John E. Schulenberg
Peter Freedman-Doan

Institute for Social Research
The University of Michigan
Ann Arbor, Michigan

2001

## TABLE OF CONTENTS

## TABLES

## FIGURES

## ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute on Drug Abuse (R01DA01411). We wish to thank Tanya Hart for her assistance with text management and editing.

## INTRODUCTION

The Monitoring the Future (MTF) project has been collecting panel survey data from several thousand adolescents (now young adults) who completed questionnaires initially as 8th grade students in 1991, 1992, and 1993. The panel surveys have been collected at two-year intervals beginning with modal age 14 and have continued through modal ages 16, 18, 20, and beyond. These panel data are a rich source of information on drug use progression, stability of drug use, long-term consequences of young adolescent substance use, links between educational success/failure and substance use, and a variety of other topics. The value of such data may be limited, however, by possible bias from two sources: (1) panel attrition, and (2) measurement error due to mode of survey administration. In this paper we describe the panel design and samples; then we explore both forms of bias, as well as ways of correcting bias, in the MTF 8th grade panel data.

The first possible source of bias we consider is sample bias due to panel attrition—the fact that some individuals targeted for follow-up surveys do not participate in further surveys. Very few potential respondents failed to participate in the 8th grade cross-sectional surveys that provided the initial (base year) data for the panels; this is fortunate because it would have been difficult to ascertain with any accuracy how non-participants differed from participants in ways that might bias findings. When non-participation occurs in follow-ups (i.e., panel attrition), the prospects for discovering and possibly correcting such biases are much better. Thus, in the case of the MTF 8th grade panels, we were able compare panel participants and non-participants in terms of characteristics measured in the 8th grade (base year) survey. For example, we explored whether panel participants (compared with non-participants) had higher grade-point averages (they did), and whether participants were more or less likely to have smoked cigarettes or consumed alcohol or used marijuana (they were less likely). Therefore, given that panel attrition was fairly severe (approximately half) by the third follow-up (modal age 20), it was clear that the panel attrition could substantially bias findings if not corrected.

Another possible source of bias in the 8th grade panel data is measurement bias due to the mode of survey administration. The initial (base year) surveys were conducted in schools under conditions of strict confidentiality. In contrast, the follow-up surveys were mailed to respondents' homes, raising the possibility that some respondents might distort certain answers (e.g., not admit the use of drugs) in order to avoid the risk that parents would discover such an admission. We were able to explore this potential source of bias by contrasting panel data with data obtained from MTF cross-sectional in-school surveys of 10th and 12th graders, which were being conducted at the same times as our mailed follow-up surveys of panel respondents. This comparison is possible because the sampling frames were very similar across the two comparison groups, indicating that any variation in responses between the groups would be due to mode of administration.

We begin by providing some background on the MTF 8th grade panel surveys and how they evolved. We next present our analysis strategy and provide further detail on the selection of the samples. Then we present our findings about both forms of bias, and our proposed means for reducing the biasing effects of panel attrition.

## Background

The MTF project has been surveying large, nationally representative cross-sectional samples of 12th grade students each year, beginning with the class of 1975. The data collections employed school-administered, self-completed questionnaires set up for optical scanning. Subsets of each senior class sample, beginning with the class of 1976, have been selected and followed up using similar questionnaires mailed at two-year intervals, thereby yielding panel data for each graduating class.

In 1991 the MTF project was expanded to include similar annual cross-sectional surveys (i.e., using large and nationally representative samples) of 8th and 10th grade students. At the time these surveys were begun, it was uncertain whether panel data would be collected from subsets of respondents; however, in anticipation of that possibility, name and address data were requested from all respondents (using procedures similar to those employed in the 12th grade surveys). Soon after the first data collection in 1991, the decision was made to collect panel data from subsets of the participants at two-year intervals.

The most important objective in conducting follow-up surveys from each cross-sectional sample of 8th grade students and 10th grade students was to obtain adequately large samples of individuals who would eventually drop out of high school, thereby compensating for the unavoidable exclusion of dropouts from the MTF 12th grade samples. This focus prompted us to oversample individuals likely to drop out before completing high school because (a) only about 15 percent of students entering high school drop out before completing, and (b) previous experience has shown that those most likely to drop out of high school are also least likely to continue to participate in panel studies. Accordingly, it was important to increase the numbers of targeted follow-up respondents who were at greater than average risk of dropping out of high school. We accomplished that objective by stratifying the 8th grade samples according to predicted likelihood of dropping out, and then giving those most likely to drop out a greater than average probability of being selected (as discussed later).

In addition to our primary purpose of obtaining nationally representative samples of high school dropouts, a second purpose for collecting panel data was to be able to study the etiology of drug use and to track changes in key attitudes and behaviors during earlier adolescence. This rationale was particularly relevant for panel data originating with 8th grade students; it would permit developmental analyses extending four years earlier than the analyses that could be conducted with MTF panel data originating with 12th grade students (i.e., 8th grade panel respondents would be tracked from modal age 14 to modal ages 16, 18, 20, and perhaps longer).

Within several years of initiating the follow-ups of 8th and 10th grade respondents, it became clear that panel attrition problems among those most at risk for dropping out were sufficiently severe that the obtained samples were judged inadequate to meet our primary initial objective of providing good samples of high school dropouts. Accordingly, we decided not to continue the initiation of new panels drawn from the 8th and 10th grade cross-sectional surveys.

By the time that decision was reached, we had already accumulated follow-up data from those originally surveyed in 1991, 1992, and 1993. We judged that the 8th grade panels for those years would be adequate to meet our second purpose—tracking development from 8th grade (modal age 14) into young adulthood. It was judged that, for the most part, this purpose did not

require a new panel to be initiated each year; the numbers of individuals targeted for follow-up across these three years totaled 6,000—a sufficiently large number of cases to permit complex analyses (even after panel attrition reduced the numbers). Accordingly, we decided to continue tracking the panel respondents initially surveyed in 8th grade in the years 1991, 1992, and 1993, while dropping further follow-ups of later 8th grade panel samples and all 10th grade panel samples. Our rationale was that we had already experienced the worst of panel attrition for the 1991-93 8th grade panels (because that occurs at the first follow-up), and the considerable investment made in collecting the early follow-up data would have greatest payoff if additional follow-up data were obtained.

**Key Issues: Representativeness and Validity of Drug Use Data**

As noted earlier, the MTF panel data obtained from adolescents initially surveyed as 8th grade students in 1991, 1992, and 1993 represent a promising source of information on drug use progression, stability of drug use, long-term consequences of young adolescent substance use, links between educational success/failure and substance use, and a variety of other topics. However, the value of these data depends, at least in part, on the answers to two sets of questions:

1. To what extent are the obtained follow-up samples *representative* of the original target samples (and the universe from which the original samples were drawn)? More specifically, does follow-up participation vary according to important individual characteristics (such as educational success, or initial levels of substance use), and does such variation appear to distort follow-up sample results? If so, to what extent can compensatory weighting (post-stratification) reduce such distortions?

2. Apart from any problems due to follow-up sample selective retention (i.e., non-random panel attrition), how *valid* are adolescents' self-reports of substance use, obtained from follow-up surveys mailed to their homes? There is evidence suggesting that adolescent respondents personally interviewed in their homes are less likely to report substance use than those surveyed in schools under conditions of anonymity or confidentiality (Harrison, 2001; Rootman & Smart, 1985). It is thus an issue of considerable interest and importance whether self-completed mail surveys sent to adolescents' homes also are subject to underreporting of substance use and other sensitive or illegal behaviors (again, in comparison with students completing the same sorts of surveys in school-based group administrations).

Two other points are worth noting here at the outset. First, it is not always the case that differential panel attrition is problematic, depending on the questions being asked of the data. In particular, while differential attrition can impact sample means and standard deviations, it often has less impact on *relations* among variables (e.g., Bryant, Schulenberg, Bachman, O'Malley, & Johnston, 2000; Schulenberg, Bachman, O'Malley, & Johnston, 1994). Thus, for research questions concerning the relational structure among variables, uncorrected differential attrition may not adversely affect conclusions. That said, the second point is that even with analyses that are primarily relational, it is clearly preferable to have the data be as accurately representative as possible. Accordingly, we consider it worthwhile to address the two sets of questions just

outlined, with the goal of finding ways to make the MTF 8th grade panel data as useful as possible.

## OVERALL ANALYSIS STRATEGY

Our broad strategy began by addressing the first set of questions above, in order to ascertain the *extent* to which retained follow-up samples differ from initial samples along key dimensions. We already knew from many earlier analyses of Monitoring the Future data (e.g., Bachman, Wadsworth, O'Malley, Johnston, & Schulenberg, 1997; Bachman, O'Malley, Schulenberg, Johnston, Bryant, & Merline, 2002; Bryant et al., 2000; Schulenberg, Bachman, O'Malley, & Johnston, 1994) that young people who have been less educationally successful and engaged are subject to greater than average panel attrition; therefore, the first step was simply to document how severe are such subgroup differences in attrition. The next step was to apply compensatory weights to reduce the distortions caused by such differential attrition. We knew also, from our own panel surveys following high school seniors into young adulthood, that individuals who used various substances as seniors are less likely than average to participate in the follow-up surveys. Compensatory weighting can be used to adjust for this sort of differential attrition as well (see, e.g., Johnston, O'Malley, & Bachman, 2001, p. 54, footnote 17); however, further weighting tailored to each dimension of substance use would be rather complex and cumbersome, and we were hopeful it would prove unnecessary.

Our analysis then turned to the second set of questions above. We compared appropriately weighted panel data from mailed surveys with data from our in-school, cross-sectional, group-administered surveys (see subsequent "Comparisons"). Lower levels of self-reported drug use in the panel data, compared with in-school survey data, could result from some degree of underreporting by some of the panel respondents—although that is not the only logical interpretation. Another possibility is that some in-school respondents exaggerate substance use, also contributing to a discrepancy. Both of these explanations involve inaccurate reporting, but another possible reason for discrepancies involves accurate reporting of *emerging* real differences between the in-school samples and retained panel samples; for example, if individuals who take up smoking after 8th grade are less likely than average to participate in mail follow-up surveys, then a "genuine" discrepancy would be evident and not attributable to any problem of inaccurate reporting by those who responded—and also not readily correctable by weighting adjustments. It follows that if the comparisons show substantial discrepancies, we cannot be sure which of these processes (and/or others) are responsible. It also follows that if no important discrepancies emerge, we cannot be sure there are not several processes in operation that tend to cancel each other. Nevertheless, the most parsimonious interpretation of a lack of substantial discrepancies would be that (a) panel attrition biases have been fairly well corrected, (b) the great majority of respondents give roughly equivalent answers whether surveyed in school or by mail, and (c) the data from the 8th grade panels can be treated as valid and roughly equivalent to data obtained from the in-school cross-sectional surveys.

To summarize, our strategy was (1) discover and (to the extent practicable, using compensatory weighting) adjust for the effects of differential panel attrition, and (2) compare the resulting adjusted panel data with cross-sectional data to see whether and to what extent there may remain evidence of substance use underreporting in the panel data.

### Comparisons to Check Representation

To explore the effects of differential panel attrition on *representation*, we focused our comparisons entirely on 8th grade data, contrasting the full *target* follow-up sample with those individuals who actually *participated* in each of the follow-ups. Thus, for example, we looked to see whether and to what extent those who were smokers in 8th grade were less likely than average to participate in the first follow-up survey, the second follow-up survey, and so on, thereby rendering the obtained samples less representative of smokers.

### Comparisons to Estimate Underreporting

For purposes of estimating *underreporting* of substance use in the mail follow-up surveys of 8th grade panel respondents, the panel data were compared with data from MTF cross-sectional in-school samples of students the same age and surveyed at the same time. Specifically, the first follow-up (which occurred two years after the initial 8th grade "base-year" survey, when respondents were modal age 16) was compared with MTF 10th grade cross-sectional surveys conducted in 1993-1995. Similarly, the second follow-up (carried out four years after the initial survey, when respondents were modal age 18) was compared with MTF 12th grade cross-sectional surveys conducted in 1995-1997.

An additional comparison was possible in which the third follow-up survey of the 1991-1993 8th graders (carried out six years later, in 1997-1999, when respondents were modal age 20) was matched with the first follow-up samples selected from MTF 12th grade respondents from the classes of 1995-1997 (specifically, those whose follow-up surveys occurred two years after graduation—i.e., in 1997-1999). This comparison, unlike those in the previous paragraph, involved no important difference in survey methods—both sets of data were obtained from mail surveys. Accordingly, because there is little reason to expect different levels of distortion such as underreporting, any remaining differences can be considered evidence of limitations in the post-stratification processes (which are necessary for both sets of data in this comparison, because both are subject to panel attrition).

If the last comparison (between those six years beyond 8th grade and those two years beyond 12th grade) shows no appreciable differences in distributions of responses to substance use items, and if the earlier comparisons (involving the first and second follow-up surveys of the 8th grade panels) do show such differences, there would be fairly strong evidence of underreporting in the first and second follow-ups of the 8th grade panels.

### Dropouts Excluded from Comparisons

For purposes of comparisons contrasting the 8th grade panel data with other samples, the 8th grade panel data exclude individuals who dropped out of school, because dropouts are not represented in the MTF in-school cross-sectional surveys.

## DESIGN OF 8th GRADE PANEL SAMPLES

The Appendix to this paper includes an extended treatment of the design and selection of the 8th grade follow-up target samples, for those wishing complete details. We present here the

broad design, focusing on those aspects necessary for an understanding of the analyses and findings reported in the remainder of this paper. Those interested in details about the initial cross-sectional surveys of 8th grade students can find that information in Bachman, Johnston, & O'Malley (2001).

**Stratification by "Academic Risk Score"**

As noted earlier, our initial purpose of obtaining sufficiently large samples of those who would later drop out of high school, as well as our realization that panel attrition would be higher among those likely to become school dropouts, prompted us to design a target sample that "oversampled" those at high risk for dropping out. We developed a risk score taking account of academic grade point average (treating those with C or lower grades as at some risk, but making no distinctions among the large majority with grades of C+ or higher), absenteeism, having been held back, having had to attend summer school, having been suspended or expelled, and parental educational attainment. The risk score had a possible range from 0 to 50, and was bracketed into four categories (0, 1-5, 6-10, 11+). Among the 8th grade respondents in the 1991-1993 cross-sectional surveys combined (and appropriately weighted to correct for initial sample variations), about 43 percent fell into the lowest risk category, 26 percent in the next to lowest, 18 percent in the next to highest, and 13 percent in the highest.

Our strategy of oversampling those at high risk of dropping out was to select individuals for the target follow-up sample such that there were 500 from each risk score stratum (for a total target of 2,000). Thus, although those at highest risk represented only 13 percent of the (appropriately weighted) cross-sectional sample, they constituted 25 percent of the follow-up target sample; conversely, those at lowest risk, although they made up 47 percent of the cross-sectional sample, also constituted only 25 percent of the follow-up target sample. Because of this sampling strategy, any particular high-risk individual in the cross-sectional sample was several times more likely than a low-risk individual to be selected for the target sample. This, of course, requires the use of corrective weights in analyses, with distinctly lower than average weights assigned to the (originally oversampled) high-risk individuals.

In sum, the 8th grade follow-up target samples heavily oversampled those who had well below average GPAs and/or had other poor school experiences (as described earlier and detailed in the Appendix). This stratification requires that corrective weights be employed in analyses, because otherwise the samples would show disproportionately large numbers of individuals with low grades and poor school experiences. Among the advantages of the stratification strategy is that it provides a counterbalancing against higher-than-average panel losses (panel attrition) among the high-risk individuals. Furthermore, the strategy has the advantage (given our research interests) of increasing the numbers (unweighted) of respondents who are involved in substance use and other problem behaviors.

**Additional Stratification Dimensions**

Although the sorting into four strata of academic risk was used primarily for oversampling those at high risk, there were a number of other dimensions of stratification included primarily to guard against accidental (random) variations that would reduce the precision of the follow-up target sample. Selection procedures were employed to make the target

samples (within each of the four risk strata) as accurately representative as possible with respect to geographical region, gender, ethnicity, and drug use; the details are included in the Appendix.

**Selection Probabilities Incorporating Adjustment for Absenteeism**

Another dimension of stratification was school absenteeism during the four weeks prior to the 8th grade in-school cross-sectional surveys. Because substance use levels are positively correlated with absenteeism, in-school surveys contain a modest degree of bias (underestimates of substance use) unless the survey data are reweighted to take account of those missing on the day (or hour) of the survey (see Johnston et al., 2001 for details). Rather than have to contend with an additional *weighting* factor in order to adjust for this problem in our analyses of the 8th grade panel samples, we opted to build such an adjustment into the selection of individuals. This was accomplished by increasing any individual's probability of being selected into the panel target sample in proportion to his or her rate of absence during the 4 weeks preceding the survey. Thus, for example, an individual who missed 3 days of school due to illness would be 17.6 percent (i.e., 20/17 = 1.176) more likely to be selected than an otherwise identical individual with no missed days (i.e., 20/20 = 1.0). An upper bound on such adjustments was set as 2.0; thus, a student who had missed more than two weeks of school during the past four weeks would have a doubled likelihood of being selected (thereby, in effect, representing both himself or herself and another student who was also ill for about two weeks during that period—including the date of the survey). Building this adjustment into the sample selection process had the advantage of avoiding any widening of the range of sample weights during the analysis phase (which would have increased error ranges in all estimates).

This adjustment for absenteeism means that the follow-up targets actually constitute slightly *less* biased samples of the initial target universes (i.e., all 8th grade students in U.S. schools) than the originally obtained cross-sectional samples (except when the latter are adjusted for absenteeism in the analysis phase). The consequence of the built-in adjustment for absenteeism, as we will see later in this paper, is that the 8th grade substance use rates for the *target panel samples*, were very slightly but consistently higher than the substance use rates for the corresponding total cross-sectional 8th grade samples from which the target panel samples were drawn.

**Questions, Analyses, and Answers About Panel Attrition and Representation**

1. *To what extent is there greater panel attrition among those who had (as of 8th grade) been relatively unsuccessful in school?* The overall answer to this question is shown by the differential attrition rates across the four original educational risk stratification categories (which were, of course, based on previous success or failure in school). More precise and refined answers could be obtained by examining various specific dimensions (GPA, etc., as reported in the base year 8th grade survey), looking to see how the obtained follow-up samples differed from the total target sample along each such dimension; however, that does not at this point seem necessary. Clearly, losses were greatest among those in the highest risk strata, as shown in the unweighted sample (topmost) portion of Table 1, and these differential effects of panel attrition were progressively more severe with each succeeding follow-up. For example, by the third follow-up (modal age 20), panel attrition had claimed only about one-quarter of

target respondents in the lowest risk stratum, compared with three-fifths of the target respondents in the highest risk stratum.

2. *To what extent is panel attrition higher among those who in 8th grade were users of cigarettes, or users of other substances?* This question was addressed by examining base year 8th grade responses, one substance at a time, contrasting those who actually participated in a given follow-up with all those originally selected and targeted for participation. Data were examined for total samples (i.e., across all four strata, taking account only of the original stratification weights) as well as for each of the strata separately; the total data provided a clear indication of the overall distortion due to panel attrition, whereas the within-strata data provided some indication of the extent to which the problem could be addressed by compensatory weighting (post-stratification, see #3 to follow). Details are included in Appendix Tables A.1.1-4 (results are shown separately for the panels initiated in 1991, 1992, and 1993, with a high degree of replication). A summary of results is provided in Table 2, showing that the retained samples do, indeed, become progressively less representative in terms of initial (i.e., 8th grade) reports of cigarette use, marijuana use (annual), and occasional heavy drinking, as well as slightly less representative in terms of alcohol use (30-day).

3. *To what extent can the distorting effects of differential panel attrition on 8th grade substance use estimates be lessened by corrective weighting (post-stratification), taking account of differential attrition among educational risk strata (see earlier question 1)?* Initial answers to this question were found by using computations similar to those in earlier question 2, except that the contributions of the four educational risk strata were (re)weighted so as to reconstitute their proportions in the original target samples. Thus, for example, individuals in the highest risk stratum were oversampled when the original target follow-up sample was drawn; therefore they were assigned relatively small weights (to compensate for their increased chance of initial selection). However, their actual rates of participation in the follow-ups were below average, so reinstating the contribution of this stratum to its original target proportion required some increase in weights. The effect of this (re)weighting was to reduce somewhat the overall variation in weights. Table 3 indicates that this initial post-stratification strategy, reconstituting the original sample proportions in terms of educational risk, eliminated the majority of the distortion shown in Table 2. In particular, the reweighted follow-up samples were quite representative in terms of 8th grade rates of smoking, alcohol use, and occasional heavy drinking; however, there remained a modest underrepresentation of 8th grade marijuana users in the obtained follow-up samples (i.e., the obtained follow-up sample, when reweighted, showed 7.6 percent annual marijuana users, in contrast to 8.3 percent in the total original sample). In balance, we judge that the post-stratification by educational risk stratum is sufficiently effective, in terms of reducing underrepresentation of 8th grade substance users, for most analysis purposes.

**Questions, Analyses, and Answers About Possible Underreporting in Follow-Ups**

1. *Are self-reports of substance use among 10th and 12th grade respondents in the first and second panel follow-ups (respectively) comparable to those of 10th and 12th grade students in the MTF cross-sectional in-school surveys at comparable years?* After properly weighting the follow-ups to correct for differential panel attrition by risk stratum, as just described, and after removing dropouts, the comparisons with the in-school cross-sections for the same years and grades/ages show (a) only very small differences for cigarette use and marijuana use (with panel follow-up data one or two percentage points lower than the corresponding cross-sectional data), and (b) slightly larger differences for 30-day alcohol use (with panel data approximately five percent lower than cross-sectional) and instances of heavy drinking (differences of less than five percent). The findings appear in Figures 1-4 and in Appendix Tables A.2.1-4. It thus appears that there may be some underreporting of substance use, especially alcohol use, among panel respondents (surveyed by mail in their homes) compared with cross-sectional respondents (surveyed in school under conditions of high confidentiality); however, it also appears that such underreporting is quite limited.

2. *Are self-reports of substance use among panel respondents at the third follow-up comparable to those of MTF seniors followed-up by mail two years* after *graduation at comparable years?* These comparisons, matching non-drop-out 8th grade panel respondents with two-year follow-up respondents from the 12th grade panels (with the 8th grade panels adjusted for differential panel attrition among risk strata and the 12th grade panels adjusted for differential attrition by initial levels of substance use), show only small and non-consistent differences between groups. The findings are included in Figures 1-4 and Appendix Tables A.2.1-4.

The first set of findings summarized above suggests to us that underreporting was slightly more frequent among 8th grade panel respondents in the first two follow-ups than among their age-mates surveyed in school. The second set of findings, that differences essentially disappeared at the third follow-up (modal age 20) when both sets of respondents were surveyed by mail, is consistent with this interpretation. It is worth adding, however, that although the age 20 data indicate that there was no *differential* underreporting by either group, these data alone cannot tell us whether that occurred because the 12th grade follow-up responses also included some underreporting, or because by age 20 there simply was very little such underreporting in either group. Obviously, both explanations can be valid to some degree, and probably are. But the more important point, in our view, is that all of the differences between groups reported here seem quite small and unlikely seriously to distort findings based on the 8th grade panel data.

**Choices and Implications of Choices for Analysis Panels**

Panel analyses involve two or more points in time, and require data from a minimum of two points in time. Ordinarily, that means that respondents who did not participate in the last data collection are deleted from the analyses, although sometimes respondents are included in analyses if they provided *any* follow-up data. Thus, for example, the analyses reported above included all respondents who provided data at whatever follow-up was being considered—that is, we did not require *complete* follow-up records in order to include respondents.

For many panel analyses, however, it is important to have complete follow-up records. This is particularly the case for complex structural equation modeling exercises involving more than two points in time.

Consistent with the above observations, there are many possible ways of defining panels for analyses. Here we focus on two:

1. A panel consisting of those respondents (from the 8th grade panels selected from among those initially surveyed in school in 1991, 1992, or 1993) who participated in *all* of the first three follow-ups (i.e., at modal ages 16, 18, and 20). This restrictive definition limits analyses to just over half (52.8 percent) of the initial target sample weighted according to the initial weights, which adjust only for the stratification according to academic risk factors.[1]

2. A panel consisting of those respondents who participated in the third follow-up (modal age 20), including some who did not participate in the first and/or second follow-up. This somewhat less restrictive definition limits analyses to 62.6 percent of the weighted cases initially targeted.[2]

For each of these two panels, it will be useful to examine drug use data at all points in time. The purpose will be to see how the findings for these restricted—but appropriately weighted—samples compare with the more complete samples. If the differences are small, then we will be comfortable allowing those retained in the panels to "stand for" those who dropped out of the panels.

*Additional Weighting Correcting also for Gender and 8th Grade GPA.* We decided that the corrective weighting for the first panel just noted should incorporate not only the four risk strata but also gender and 8th grade grade point average (GPA).[3] The reasons for these additional weighting adjustments are outlined in the following paragraphs.

We decided that our re-weighting (restratification) should take account of gender as well as risk stratum, for the following reasons: (a) we consider it likely that many 8th grade panel analyses will be carried out separately for males and females; (b) panel attrition is higher among males than among females; therefore (c) analyses for separate genders based on a re-weighting that did *not* incorporate gender would involve distortions.

---

[1] The 52.8 percent value is the most meaningful *response rate*, because it corrects for the fact that the initial sample included a disproportionately high number of "high risk" cases. The corresponding unweighted response rate (i.e., the raw proportion of target respondents actually obtained) is 46.5 percent. See Appendix Table A.3 for further details.

[2] The 62.6 percent value is the most meaningful *response rate*, for reasons explained in the previous note. The corresponding unweighted response rate is 56.7 percent (see also Appendix Table A-3).

[3] Out of the original target surplus totaling 6,000 cases, 163 (2.7 percent) had missing data on the gender variable and were excluded from the sample in this weighting scheme. Missing data on 8th grade GPA was simply treated as another category of GPA; thus, no additional cases were excluded.

Our decision to include 8th grade GPA as a further basis of stratification was reached after discovering that failure to do so resulted in some degree of distortion in the retained sample. Specifically, the retained sample showed higher 8th grade GPA, on average, than did the original target sample. As was the case for substance use, corrective weighting to reconstitute proportions in the original risk strata reduced this distortion considerably; however, there remained a slight bias upward in GPA within each stratum (and thus also for the total), as we illustrate in Table 4 and discuss in the next paragraph. We undertook the further post-stratification by 8th grade GPA because we anticipated future analyses would focus extensively on aspects of educational attainment, and therefore considered it useful to have a weighting that would make the retained sample as representative as possible along that dimension. We felt confident that such further re-weighting would not make the sample any less representative in terms of substance use, and that proved to be correct. We also hoped that it might actually make the sample a bit more representative in terms of substance use, and it did—but only to a very slight degree, as we show later in Tables 5-8.

Table 4 illustrates how these different re-weighting schemes operate on a variety of sub-samples of the panel with respect to weighted numbers of cases and mean levels of GPA. The first four rows of entries (1-4) consist of panels of respondents who participated in the original base year (8th grade target sample) survey (row 1) and in at least one subsequent follow-up survey (rows 2-4). The means and numbers of cases displayed in these rows have been re-weighted to adjust for the initial probability of selection by risk stratum and by differential panel attrition by risk stratum. Rows 5a, 5b, and 5c are panels consisting of respondents who participated in the original base year survey and in *all three* subsequent follow-up surveys. As the row headings indicate, results in row 5a were re-weighted to adjust for the initial differential probability of selection by risk stratum and by differential panel attrition by risk stratum. Results in row 5b show the effects of adding gender to the re-weighting scheme, and row 5c adds both gender and 8th grade self-reported GPA.

The first column of Table 4 shows how panel attrition, when left uncorrected, gradually distorts 8th grade GPA. The mean self-reported GPA for the target sample was 5.73. With no re-weighting for differential panel attrition, the subsample consisting of those who participated at the first follow-up had a mean GPA of 5.95. When the panel is restricted to those who participated in all three follow-ups the mean, unadjusted GPA rises to 6.27.[4]

The second column in Table 4 also shows 8th grade GPA, but this time with weights adjusted to take account of differential panel attrition. In this column the distortion is much smaller for all the subsamples in the table. In the subsample of participants in all three follow-ups (Rows 5a, 5b, and 5c), the re-weighting to account for differential attrition by risk stratum (Row 5a) eliminates a great deal of the bias due to panel attrition (i.e., a mean of 5.93 versus the unadjusted value of 6.27). Adding gender and GPA to the re-weighting eliminates virtually all of the distortion due to panel attrition (i.e., a mean of 5.76 versus the original target mean of 5.73).

[4] On the scale used, a "5" refers to a GPA of B-, a "6" refers to a B, and a "7" refers to a B+. The standard deviation on the 8th grade GPA scale is approximately 2.2.

The remaining columns in Table 4 show GPAs for 10th grade and 12th grade, and the bottom row (5c, the only row that includes stratification according to original GPA reports) comes closest to the top entry—i.e., does the "best" job of approximating the data from the largest available sample. Our conclusion, based on this analysis of the data in Table 4, is as follows: the final weighting scheme, when applied to just that portion of the sample that participated in all three follow-up surveys, does a very good job of reproducing the GPA distribution for the largest available sample.[5]

Tables 5-8 are parallel in design to Table 4, and show how the several different weighting schemes worked for a number of substance use dimensions at modal ages 14, 16, 18, and 20. Table 5 shows that the final weighting including the GPA adjustment (row 5c) produces estimates that are consistently slightly higher than the weightings without GPA adjustment (rows 5a and 5b); the table shows also that the bottom row estimates are generally quite close to the estimates with the least constrained sample (i.e., those at the top of each column). Tables 6 and 7 show much the same pattern for the two measures of alcohol use. Table 8 shows that even after extensive adjustment, panel attrition reduces the estimates of marijuana use among 8th grade students (modal age 14) in the retained sample; however, the follow-up estimates (modal ages 16, 18, and 20) all reproduce quite closely the estimates based on the (weighted) full follow-up samples available at each follow-up.

Based on these comparisons, we conclude that the 8th grade panel sample restricted to those who participated in all three of the follow-ups, and weighted according to our final stratification scheme (which post-stratifies according to original risk stratum, gender, and 8th grade GPA), does quite well in representing the characteristics of the larger total available samples at 8th grade and at each of the follow-ups; these well-represented characteristics include cigarette use (see Table 5), alcohol use (see Table 6), occasional heavy drinking (see Table 7), marijuana use (see Table 8), and grade point average (see Table 4).

Appendix Table A.4 includes a full listing of the weights used in the final stratification scheme, including numbers of respondents to which each weight is assigned. A summary of the weights and weighting scheme (collapsed across the three cohorts—1991, 1992, and 1993) is presented in Table 9. In general, these weights vary less widely than the original set of weights (before any adjustment for differential panel attrition). As can be seen in Table 9, the weights for female respondents are, with only one small exception, consistently smaller than the corresponding weights (i.e., matched in terms of risk stratum and GPA) for males. This corrects for the fact, discussed earlier, that response rates were distinctly higher for females than for males; moreover, the table illustrates that this gender difference in panel attrition was evident at virtually every level of academic risk and GPA. In other words, even after we adjust for male-female differences in academic success (in 8th grade), there still are substantial gender

[5] It should be kept in mind that the 8th grade (modal age 14) responses shown in Table 4 were obtained in the initial (base year) data collection and refer to 8th grade GPA, data from the first follow-up (modal age 16) refer to 10th grade GPA, and data from the second follow-up (modal age 18) refer to 12th grade GPA. A very small number of respondents in the first follow-up, and a modest number of respondents in the second follow-up, had dropped out of school; such individuals were reporting their most recent GPAs, not necessarily 10th or 12th grade GPAs. These few exceptions do not interfere with the point being illustrated in Table 4, and thus, we saw no need to screen them out of the Table 4 analyses.

differences in response rates that require additional corrective weighting in order for the obtained panel sample to be optimally representative.

**Design Effects**

The Monitoring the Future surveys of school students employ complex sampling strategies that include stratification, clustering by schools, and differential weighting of respondent scores (i.e., some respondents are given more weight than others). Consequently, statistical tests designed for simple random samples may not be entirely accurate. This is because stratification tends to heighten the precision of a sample (compared with a simple random sample), whereas clustering and weighting tend to reduce precision. The Monitoring the Future samples of students in schools are appreciably less efficient than simple random samples, particularly because of the clustering and differential weighting involved. The smaller follow-up panel samples, however, are not so much less efficient than simple random samples, because clustering and differential weighting are much less than in the large in-school samples.

Various methods now exist for estimating the degree to which a complex sample departs from a simple random sample in terms of its precision. A measure of this departure was described by Kish (1965) as the *design effect*, which he defined as the actual sampling variance (from a complex sample) divided by the expected sampling variance from a simple random sample with the same number of elements. Thus, for example, a design effect of 2.0 would mean that a complex sample of 2,000 cases would have precision equivalent to a simple random sample of 1,000. Clearly, the smaller the design effect, other things equal, the better in terms of precision.

We calculated design effects for a very limited number of statistics in order to compare the several sample/weighting choices possible for the 8th grade panel data.[6] Table 10 shows design effects (along with numbers of weighted cases) for a panel sample limited to participants in the third follow-up (which includes some individuals who missed the first and/or second follow-up). Table 11 provides comparable data limited to participants in *all three* follow-ups. The weighting scheme in Table 11 takes account of gender and 8th grade GPA, as well as differential attrition by risk stratum (discussed earlier), whereas the weighting used in Table 10 corrects only for differential attrition by risk stratum. A number of observations can be made based on examination and comparison of Tables 10 and 11:

1. The numbers of cases are, of course, different. There are fewer cases for each entry in Table 11 than for the corresponding entry in Table 10. However, it can also be seen that the numbers across follow-ups are nearly identical in Table 11 (differing only due to small amounts of missing data on specific drug use measures), whereas in Table 10 the numbers of cases are distinctly smaller for the first and second follow-ups than for the third.

[6] Design effects were estimated using IVEware software, developed at the University of Michigan in 1998 by T. E. Raghunathan, Peter W. Solenberger, and John Van Hoewyk. IVEware calculates sample variances using the Taylor series expansion method.

2. The mean scores, which reflect proportions of users for each of four substance use measures, are very similar across the two tables. This is reassuring, because it suggests that if we choose to use the more restricted sample (as in Table 11), the elimination of some respondents because of missing data on one (or both) of the earlier follow-ups does not substantially distort the substance use estimates.

3. The design effects for the more restricted sample in Table 11 are generally smaller than those in Table 10. More important, this means that the *effective Ns* in Table 11 are not much smaller than those in Table 10; indeed, in 6 out of 16 comparisons, it appears that the more restricted sample is actually more precise (i.e., the effective N is larger) compared with the less restricted sample.

All of this leaves us confident that when we conduct analyses restricted to the limited set of respondents who participated in all three follow-ups, the reduced numbers of cases (compared with a less restrictive choice) costs us very little in terms of accuracy.

## SUMMARY AND CONCLUSIONS

This paper describes the Monitoring the Future panel data from individuals initially surveyed as 8th grade students in 1991, 1992, and 1993. More important, the paper examines two possible sources of bias in these data: bias due to panel attrition, and bias in substance use reports due to mode of survey administration.

### Stratified Follow-Up Target Samples

The samples targeted for follow-up consisted of 2,000 students in each of the three 8th grade surveys (1991, 1992, and 1993), for a total of 6,000. The follow-up target samples were stratified by "academic risk score" such that individuals at greatest risk for dropping out of high school were oversampled. The individuals at greatest risk of dropping out of high school were also the ones we were least likely to locate successfully and the ones least likely to respond, even if we located them. Thus, this disproportionate sampling served as somewhat of a hedge against panel attrition—in that the strata most likely to suffer high losses in follow-up surveys were overrepresented from the start.

### Biases Due to Panel Attrition

As expected, we found that panel losses (inability to obtain follow-up responses) had the effect of biasing the obtained samples. Specifically, the obtained follow-up samples were more likely than the original (target) samples to include (a) individuals with little or no substance use (along various dimensions, as reported in the 8th grade survey), (b) those with above average 8th grade GPAs (grade point averages), and (c) females. These biases grew larger in the later follow-ups, with their proportionately larger panel losses.

Fortunately, we found that post-stratification—the re-weighting of sample participants so as to reconstitute the original sample distributions in terms of academic risk (and also gender and GPA, in later analyses)—had the effect of greatly reducing the biases in substance use measures. When the re-weighting included gender and GPA, the panel attrition biases along those

dimensions were virtually eliminated. In other words, our adjustments of the weights for obtained follow-up sample participants had the effect of giving slightly increased weight to males, those with lower GPAs, and those in the "high risk" strata, the effect of which was to render the reweighted obtained samples much more "accurate" or "correctly balanced" in terms of 8th grade reports of substance use, as well as gender and 8th grade GPA. An additional effect of the re-weighting was to make the follow-up samples more efficient, because it gave somewhat increased weights to individuals who had originally been oversampled and thus had relatively small weights. *Our conclusion from this aspect of the analysis is that the post-stratification re-weighting is highly useful and appropriate; it makes the retained sample much more accurately representative of the original cross-sections of 8th grade students (and, thus, the target universe of 8th graders).*

**Biases in Follow-Up Substance Use Reports**

We were concerned that the follow-up surveys mailed to the homes of respondents, especially for the first two follow-ups (occurring at modal ages 16 and 18), might inhibit some individuals from reporting substance use. When we compared their results with those from in-school surveys of those the same age (i.e., when 10th and 12th grade students were compared with appropriately weighted respondents to the first two follow-ups, respectively, with comparisons restricted to those still in school), we found only very modest differences in percentages reporting smoking, alcohol use, occasional heavy drinking, and marijuana use. *We thus conclude that underreporting biases due to mail follow-up procedures are generally quite small, and that they will not seriously distort analyses using the 8th grade panel data—provided, of course, that the re-weighting described previously is used.*

**Conclusion**

The post-stratification re-weighting approach outlined and illustrated here seems to us to be highly appropriate and entirely free of "downside risks." It takes data which are already weighted—necessarily so—and revises the weights so as to reproduce the original sample more accurately. Moreover, the new weighting scheme shows a somewhat narrower distribution of weights, with the result that the sample is more "efficient" than if the original weights were used. Most importantly, the sample estimates resulting from the new weighting scheme are quite realistic in terms of substance use, as well as educational success and gender. We wish all analysis decisions were as easy as this one!

## REFERENCES

Bachman, J. G., Johnston, L. D., & O'Malley, P. M. (2001). *The Monitoring the Future project after 27 years: Design and procedures.* (Monitoring the Future Occasional Paper No. 54). Ann Arbor, MI: Institute for Social Research.

Bachman, J. G., O'Malley, P. M., Schulenberg, J. E., Johnston, L. D., Bryant, A. L., & Merline, A. C. (2002). *The decline of substance use in young adulthood: Changes in social activities, roles, and beliefs.* Mahwah, NJ: Lawrence Erlbaum Associates.

Bachman, J. G., Wadsworth, K. N., O'Malley, P. M., Johnston, L. D., & Schulenberg, J. (1997). *Smoking, drinking, and drug use in young adulthood: The impacts of new freedoms and new responsibilities.* Mahwah, NJ: Lawrence Erlbaum Associates.

Bryant, A. L., Schulenberg, J., Bachman, J. G., O'Malley, P. M., & Johnston, L. D. (2000). Understanding the links among school misbehavior, academic achievement, and cigarette use: A national panel study of adolescents. *Prevention Science, 1*(2), 71-87.

Harrison, L. (2001). Understanding the differences in youth drug prevalence rates produced by the MTF, NHSDA, and YRBS studies. *Journal of Drug Issues, 31*, 665-694.

Johnston, L. D., O'Malley, P. M., & Bachman, J. G. (2000a). *Monitoring the Future national survey results on drug use, 1975-1999. Volume I: Secondary school students*. (NIH Pub. No. 00-4802). *Volume II: College students and adults ages 19-40*. (NIH Pub. No. 00-4803). Bethesda, MD: National Institute on Drug Abuse.

Kish, L. (1965). *Survey sampling.* New York: John Wiley & Sons.

Rootman, I., & Smart, R.G. (1985). *A comparison of alcohol, tobacco and drug use as determined from household and school surveys.* Drug and Alcohol Dependence, 16, 89-94.

Schulenberg, J., Bachman, J. G., O'Malley, P. M., & Johnston, L. D. (1994). High school educational success and subsequent substance use: A panel analysis following adolescents in young adulthood. *Journal of Health and Social Behavior 35*(1), 45-62.

# TABLES

# FIGURES

# APPENDIX

This Appendix contains 1) a series of tables that provide full details of results summarized in the body of the Occasional Paper, and 2) a detailed description of the design and selection procedures utilized for the 8th grade follow-up samples.

## Appendix Tables

Tables A.1.1-4 show, separately by substance and year, the effects of panel attrition on substance use in the panel samples. These tables provide details for results summarized in Table 2 of the main text. Tables A.2.1a-A.2.4c present comparisons of substance use (by year) between the 8th grade panels and the comparable MTF cross section. These tables provide details for results depicted in Figures 1-4. Table A.3a reports on panel attrition and results of adjustments for panel attrition in a subsample composed of respondents to all three follow-ups. Table A.3b reports on panel attrition and results of adjustments for panel attrition by gender. It is useful to compare these tables to Table 1 in the text. Tables A.4a-c supply full details, by year, of the weights used in the final stratification scheme. Table 9 in the text summarizes these details.

## Design and Selection of 8th Grade Follow-up Samples

This section outlines the procedures used to select follow-up respondents from the Monitoring the Future (MTF) base year (BY) samples of 8th graders. The procedures described were used to select follow-up respondents from the 8th grade class of base years 1991, 1992, and 1993.

Our underlying design objective was to generate annual panels of young people that would overrepresent respondents who would fail to earn a high school diploma in the usual time frame. Our working definition of "the usual time frame" was four complete school years after being enrolled in 8th grade.

Our sampling objective was to select 2,000 cases from each 8th grade base year cohort. The frame for the 8th grade panels included all 8th graders who gave us usable data in the base year data collection. Base year respondents who did not provide us with a valid *name* on the name and address card were defined as ineligible and were removed from the frame. We also removed from the frame any respondent with missing data on all components of the risk score index (see point 1 of the following paragraph). There were, therefore, three criteria for determining eligibility for this frame: usable 8th grade data, valid *name* on the tear-off card, and valid data on at least one item in the risk score index.

Stratification variables for the panels included risk score, region, base year school ID, questionnaire form, sex, ethnicity, drug use, and random individual ID. These variables were used in the order listed.

A. **Risk Score**: Each BY respondent was assigned a risk score on an index of educational success. The risk score was the mean computed by 1) averaging the respondent's answers to seven variables, and then 2) multiplying by 10 (and rounding to integer). The mean was computed using however many valid answers the respondent provided, as long as the

respondent provided at least one valid answer. The variables' response categories and recoded values that make up the index are displayed in Table A.5. The lowest score possible on this index is 0, while the highest score possible, with no missing data, is 33. The highest score would result from summing the highest recoded score on each variable (a total of 23), dividing by seven (23/7 = 3.3), and multiplying by 10 (3.3 x 10 = 33). For cases with missing data, the highest possible score is 50, because a code of 50 on the only non-missing variable would simply be multiplied by 10. The highest scores actually obtained in a test run were in the high 30s and low 40s. Our selection strategy called for selecting follow-up respondents according to the distribution displayed in Table A.6.

B. **Region**: Used the four standard regions of the U.S.:

North East, North Central, South, and West

C. **Base Year School ID**: A constant for everyone from the same school, to keep schools together for sampling purposes. No distinction is made between public and private schools.

D. **Questionnaire Form**: 8th/10th grade base year questionnaire form 1 or 2

E. **Sex**: Male, Female, Unknown

F. **Ethnicity**: Used the fully detailed variable

G. **Drug Use**: A five category index variable:

a. Never used cigarettes, alcohol, or other drugs (lifetime). Category includes cases with missing data on all component variables.

b. Used cigarettes or alcohol (lifetime), but not in the past 30 days. Never used other drugs (lifetime).

c. Used cigarettes or alcohol in the past 30 days. Never used other drugs (lifetime).

d. Used marijuana (lifetime) or used other drugs on 1 or 2 occasions (lifetime).

e. Used other drugs on 3 or more occasions (lifetime) or any use of heroin (lifetime).

Component variables recoded to construct the drug use index: lifetime cigarette use dichotomy, 30-day cigarette use dichotomy, lifetime alcohol use dichotomy, 30-day alcohol use dichotomy, lifetime (illicit) drug use index.

H. **Random Individual ID**: A randomly assigned 10-digit ID

Sample selection required the following four steps, carried out separately within each risk score stratum to provide strict control of the sample size.

1. Calculate a selection weight for each individual using the following formula. The formula adjusts the school-level base year selection weight for individual absenteeism as reported in

the base year questionnaire. We used a modified version of the absenteeism weight traditionally used for 12th graders.

$$\textit{Selection Weight} = \frac{\textit{Base Year School Weight} * 20}{20 - \textit{Number of Days Absent}}$$

The sampling section provided MTF with 8th grade school weights already adjusted for within-school response rates; however MTF continues to use the "traditional school weights" *not* adjusted for within-school response rates because the sampling's response rate adjustments were based on pre-cleaned and aggregated data. The base year school weight used here is the traditional school weight, since we already incorporated an explicit measure of individual absenteeism in this step.

The number of days absent was calculated from the responses given to the base year questions:

During the LAST FOUR WEEKS how many whole days of school have you missed
a. Because of illness
b. Because you skipped or "cut"
c. For other reasons

The responses to these three questions were recoded as follows, imputing the mean for any missing data:

| | | |
|---|---|---|
| 1.0 Day | = | 0.0 |
| 2.1 Days | = | 1.0 |
| 3.2 Days | = | 2.0 |
| 4.3 Days | = | 3.0 |
| 5.4 or 5 Days | = | 4.3 |
| 6.6 to 10 Days | = | 7.2 |
| 7.11 or more Days | = | 10.0 |

The recoded (or imputed) values for each question were then summed to form a Number of Days Absent score for each respondent. This sum variable has a maximum value of 10, with any values greater than 10 being recoded to 10.

This measure differs from the one used with the 12th grade MTF panels in three ways:

(1) The mean was imputed for missing data instead of setting missing data to 0 days.
(2) Impossibly high levels of absenteeism were set to the maximum of 10 days instead of 0 days.

(3) The maximum was set at 10 days instead of 16 days, ensuring that no one's selection weight was more than double his or her base year school weight.

2. Calculate a cumulative sum of the selection weights and attach the "running cumulative sum" to each of the sorted data records. The last data record for each stratum ended up with the total sum of selection weights for that stratum.

3. Calculate the selection interval within each stratum by dividing each stratum's total sum of selection weights by the target sample size for that stratum.

4. Take interval selections from the sorted lists using the calculated intervals. A case was selected into the sample whenever the cumulative sum of selection weights attached to its record exceeded the incremented interval. The interval was incremented only when a selection was made. For example, as shown in Table A.7, in a stratum with an interval of 4.09, the cases attached to cumulative sums exceeding 4.09, 8.18, 12.27, 16.36, 20.45, 24.54, 28.63, 32.72, 36.81, . . . would be selected into the sample. We provided for a random start by using random individual ID as the last sort variable.

Whenever a case had a selection weight larger than the selection interval for the stratum, it caused subsequent, adjacent cases to come into the sample even though these cases may have had small selection weights. This strategy makes particular sense when working with a carefully stratified frame since adjacent cases are then very similar on all key variables. Table A.7 illustrates what might happen using the example numbers.

**Table A.5**

Risk Score Recoding

| Variable | Original Response | Original Code | Recode |
|---|---|---|---|
| GPA | C + or higher | 4-9 | 0 |
| (this school year) | C | 3 | 1 |
| | C- | 2 | 2 |
| | D | 1 | 3 |
| Days Skipped | None | 1 | 0 |
| (in last 4 weeks) | 1 | 2 | 1 |
| | 2 or more | 3-7 | 2 |
| Classes Skipped | None | 1 | 0 |
| (in last 4 weeks) | 1-2 | 2 | 1 |
| | 3 or more | 3-6 | 2 |
| Held Back | Never | 1 | 0 |
| (ever) | Once | 2 | 3 |
| | Two or more times | 3 | 5 |
| Summer School | Never | 1 | 0 |
| (ever) | Once | 2 | 1 |
| | Two or more times | 3-4 | 2 |
| Suspended/Expelled | Never | 1 | 0 |
| (ever) | Once | 2 | 3 |
| | Two or more times | 3 | 5 |
| AveParentEduc | 30 or more (both HS grads) | ≥ 30 | 0 |
| | 25 | 25 | 1 |
| | 20 | 20 | 2 |
| | 15 | 15 | 3 |
| | 10 (both grade school) | 10 | 4 |

**Table A.6**

Distribution of Risk Strata in 8th Grade (Class Year 1991)

| RISK SCORE VALUES | | | | | |
|---|---|---|---|---|---|
| | 0 | 1-5 | 6-10 | 11+ | Total |
| N selected | 500 | 500 | 500 | 500 | 2000 |
| BY91 N | 8131 | 4606 | 2915 | 1609 | 17262 |
| Percentage | 47% | 27% | 17% | 9% | 100% |
| Sample rate | .06 | .11 | .17 | .31 | |

**Table A.7**

Example of Sample Selection Within the High Risk Stratum

| Case # | Selection Weight | Cumulative Sum of Weights | Incremented Interval | Cases Selected |
|---|---|---|---|---|
| 1 | 1.5 | 1.50 | 4.09 | |
| 2 | 1.2 | 2.70 | 4.09 | |
| 3 | 1.4 | 4.10 | 4.09 | X |
| 4 | 0.7 | 4.80 | 8.18 | |
| 5 | 0.5 | 5.30 | 8.18 | |
| 6 | 1.0 | 6.30 | 8.18 | |
| 7 | 2.0 | 8.30 | 8.18 | X |
| 8 | 1.3 | 9.60 | 12.27 | |
| 9 | 0.8 | 10.40 | 12.27 | |
| 10 | 8.8 | 19.20 | 12.27 | X |
| 11 | 3.1 | 22.30 | 16.36 | X |
| 12 | 1.9 | 24.20 | 20.45 | X |
| 13 | 1.3 | 25.50 | 24.54 | X |
| 14 | 1.1 | 26.60 | 28.63 | |
| 15 | 0.9 | 27.50 | 28.63 | X |
| 16 | 2.1 | 29.60 | 28.63 | |
| 17 | 1.1 | 30.70 | 32.72 | |
| 18 | 0.6 | 31.30 | 32.72 | |
| 19 | 0.3 | 31.60 | 32.72 | |
| 20 | 0.9 | 32.50 | 32.72 | |