Appendix A
Sample Design, Development of Weights, Confidence Intervals, and Data Suppression

This appendix provides a detailed discussion of the sample design and estimation procedures employed for the NSPY. Section A.1 describes the multistage stratified sample design used to select youth and their parents for the evaluation. Section A.2 describes the procedures used to construct sampling weights and related replicate weights for analysis and calculation of sampling errors. Section A.3 summarizes the procedures used to calculate confidence intervals, rules used for data suppression, and design effects for selected statistics.

A.1 Overview of Sample Design

Youth and their parents were found by door-to-door screening of a scientifically selected sample of 81,000 dwelling units, of which 34,700 dwelling units were screened in Wave 1, 23,000 dwelling units were screened in Wave 2, and 23,300 dwelling units were screened in Wave 3. The sampled dwelling units were spread across about 2,900 neighborhoods (i.e., area and building permit segments defined below) in a diverse set of 90 large geographic areas referred to as primary sampling units (PSUs). The sample was selected in such a manner as to provide an efficient and nearly unbiased cross-section of America's youth and their parents. All types of residential dwelling units were included in the sample. However, institutions, group homes, and dormitories were excluded from the study.

The selection of dwelling units and youth was done only for the three initial recruitment waves. For subsequent followup waves (i.e., Waves 4 through 9), there was no new selection of dwelling units or youth. However, an originally-sampled parent could be replaced by a new parent under certain conditions described below.

The sampling of youth was designed to obtain sufficient numbers of youth in each of three targeted age ranges: 9 to 11 years, 12 to 13 years, and 14 to 18 years. These age ranges were judged to be important analytically for evaluating the impact of the Campaign. Within households with multiple eligible youth, up to two youth were selected for the study.

Parents were defined to include natural parents, adoptive parents, foster parents, or other caregivers who lived in the same household as the sample youth. Stepparents were eligible for the study unless they had lived with the child for less than 6 months. If more than one parent or caregiver was present in the household, one was randomly selected with no preference given to selecting mothers over fathers. In the followup waves, if the original parent was no longer eligible (e.g., no longer lived with child for at least 2 nights a week, or was mentally or physically disabled) another parent who was considered most knowledgeable about the sampled youth was selected as a replacement. When two youth who were not siblings were selected for the study (i.e., in multifamily households), a parent/caregiver was selected for each.

The following discussion is divided into four subsections. The first two subsections describe the selection of the screening and eligible household samples. The remaining two subsections describe the selection of youth and parents for the recruitment and followup waves. As mentioned earlier, all of the major sampling activities occurred during the three initial recruitment waves. The samples for the followup waves were generally subsets of youth and parents selected for a prior wave.

A.1.1 Selection of Screening Sample (Waves 1 through 3)

The screening sample was selected using a dual-frame multistage design. The primary sampling frame consisted of lists of dwelling units built by late 1991 that had been compiled by Westat for a sample of geographic areas called "area segments." This frame is referred to as the area frame. The second frame consisted of building permits issued between January 1990 and December 1998 for new housing units. This frame is called the building permit frame. By sampling new construction from the building permit frame, it was possible to distribute the sample more effectively within PSUs and thus achieve improved sampling precision (Judkins, Cadell, and Sczerba, 2000). Dwelling units built in 1990 and 1991 had two chances of selection since they could appear in both frames. To account for possible duplication, the screening questionnaire in Waves 1 through 3 included a question on the age of the dwelling unit. Any dwelling unit in the area frame built after April 1, 1990 was considered to be ineligible for the sample since it could have been listed in the building permit frame. On the other hand, any dwelling unit built during the first 3 months of 1990 was retained in the area sample under the assumption that there is normally a lag between the issuance of the building permit and the construction of the building.

There are some coverage limitations in the procedures used to select the dwelling units. Dwelling units built in 1999 or later had no chance of selection from either frame. Also, dwelling units built during the 1990s in jurisdictions where no permit was required had no chance of selection. Finally, modular housing units built during the 1990s were not included in the building permit frame and thus had no chance of selection. These three factors implied a dwelling unit coverage rate of about 98 percent.

A.1.1.1 Selection of the Area Sample (Waves 1 through 3)

The area sample was selected in three stages. The first stage consisted of selecting a sample of PSUs. The PSUs were generally metropolitan statistical areas (MSAs) or groups of nonmetropolitan counties. The second stage consisted of selecting area segments within PSUs. An area segment was defined to be a Census-defined block or group of contiguous blocks with a minimum dwelling unit count of 60, based on the 1990 Census of Population. The third stage consisted of selecting dwelling units within the sampled area segments. Details about the three stages of selection are given below.

PSU Selection

The primary sampling units (PSUs) were generally defined to be metropolitan statistical areas or groups of rural counties. The PSUs defined for sampling purposes were constructed using 1990 Decennial Census information and met the following general criteria:

- Each PSU consisted of a single county, a group of counties, or a metropolitan statistical area (MSA).

- No PSU had a 1990 population larger than 5,400,000. (In order to meet this criterion, the New York, Chicago, and Los Angeles metropolitan areas were divided into three, two, and two PSUs, respectively.)
- The PSUs were geographically contiguous, mutually exclusive, and covered the United States.
- Nonmetropolitan PSUs did not cross state boundaries.
- Each PSU had a total population of at least 15,000 as of the 1990 Census.
- Each PSU was designed to be easily traversable by an interviewer or lister, given population density, minimum size constraints, and natural topography.

The PSU sampling frame included a total of 1,404 PSUs, from which a sample of 100 PSUs had originally been selected for a previous study as follows. First, the 24 PSUs with populations greater than 2,100,000 were made "certainty" selections (i.e., selected with probability 1). The 24 certainties included three PSUs in New York, two PSUs in Chicago, and two PSUs in Los Angeles. The remaining 1,380 noncertainty PSUs were then assigned to 38 strata defined by region, metropolitan status, per capita income, percent minority population, and population size of PSU. Two PSUs were randomly selected from each stratum with probabilities proportionate to 1990 population using the Durbin-Brewer method (Durbin, 1967).

For the NSPY, a subsample of 90 PSUs was selected from the original 100-PSU sample. An important reason for using a subset of the 100 PSUs instead of selecting a fresh set of 90 PSUs was that Westat had experienced interviewers in these PSUs. In addition, it was possible to use area listings from the prior survey, thereby reducing the area sampling costs. To select the subsample, the noncertainty strata and two pairs of small certainty PSUs were grouped into 10 superstrata consisting of four original strata per superstratum. One stratum was then randomly selected from each superstratum. Within the selected stratum, one of the two previously-selected PSUs was randomly deleted. This left a total of 90 PSUs for the NSPY. Additional details about the PSU selection process are given in Rizzo and Judkins (2004).

Area Segment Selection

Area segments consisted of groups of neighboring blocks with a minimum count of 60 dwelling units based on 1990 Census information. By using blocks instead of larger units of geography, such as Census-defined tracts or block groups, the size of the listing task could be reduced. Blocks with very small (or zero) numbers of dwelling units were collapsed with adjacent blocks to meet the minimum size requirement. The use of a relatively large minimum size of 60 dwelling units helped avoid the selection of neighboring households for the sample. A total of 1,180 area segments was selected for Wave 1, while 689 segments were selected for Wave 2 and 694 segments were selected for Wave 3.

The segments selected for Wave 1 consisted of a subset of segments that had been selected for another survey conducted in late 1991. Since listing is a relatively costly operation that requires sending field workers to each sampled segment to prepare lists of dwelling units, the use of the existing sample saved the costs associated with a complete relisting of 1,180 area segments. The use of the old listings also had the advantage of not including most dwelling units built during the 1990s. While the lack of coverage of 1990s construction would ordinarily be a drawback, the sample design for the NSPY included a separate stratum of building permits to cover 1990s construction. Thus, using an old list actually made the screening effort more efficient because any dwelling units built in the 1990s in these

area segments would be screened out of the sample. On the other hand, the area segments for Waves 2 and 3 were completely relisted in the fall of 1999.

A fixed whole number of segments was allocated to each PSU based on the projected count of 9- to 18-year-olds in the stratum containing the PSU. A total of 2,065 segments was available from the earlier study. The segments had been selected systematically with probabilities proportionate to size (PPS)[1] where the measure of size was a weighted measure that gave African American and Hispanic households greater weight than other households. The use of a weighted measure of size was desirable for the earlier study and resulted in an oversample of segments with high concentrations of minority populations. However, this oversample was not efficient for the NSPY. Since only 1,180 segments were required, a subsample of segments was drawn in a way that resulted in overall probabilities of selection that were proportional to total household counts without any special emphasis on minority households. This was accomplished by using a sampling measure of size (SEGMOS) that was proportional to the ratio of desired overall probability to the original probability:

$$SEGMOS = \frac{\text{Number of households in segment in 1990}}{\text{MOS of segment for original survey}}.$$

Dwelling Unit Selection in Area Segments

For Wave 1, a systematic PPS sample of 30,993 dwelling units was drawn from the 1,180 area segments that had been listed in late 1991 and early 1992. When combined with the permit sample of 3,407 newly built dwelling units described below, the total initial sample size in Wave 1 was 34,400. Within the selected area segments, dwelling units were subsampled at rates designed to yield an overall equi-probability sample of dwelling units (i.e., a sample in which every dwelling unit had the same overall chance of selection). For Wave 2, the sample of dwelling units in the 689 area segments was supplemented with 2,875 dwelling units from the building permit frame for a total of 23,000 dwelling units. For Wave 3, the sample of dwelling units in the 694 area segments was supplemented with 3,052 dwelling units from the building permit frame for a total of 23,300 dwelling units.

For a subsample of the selected dwelling units, there was a quality control check on the original 1991-92 listing procedures. For all single-family dwelling units, the interviewer checked for hidden apartments (such as converted basements, garages, and attics) that might have been missed by the original lister. Any hidden apartments identified by this process were added to the sample. Also, in a subsample of multifamily dwelling structures, the interviewer checked for missed apartments. As a result of these quality control procedures, 192 missed dwelling units were added to the sample in Wave 1.

New mobile homes placed on sites between 1991 and 2000 had a chance of selection through the missed mobile home procedure. In a sample of area segments, interviewers were instructed to locate and enumerate mobile homes on their first visit and to compare what they found with what was found when the segment was first listed in 1991. In this sample of segments, any new mobile homes found were added to the sample. If there were more than nine new mobile homes in a segment (as might be the case with a new mobile home park), a subsample was drawn and appropriately weighted. This procedure added 99 mobile homes to the sample. Thus the combined sample from area segments was

[1] Systematic PPS (probability proportionate to size) selection refers to a method of selection in which the sampling frame is sorted in a prescribed manner and a systematic sample is drawn from the sorted frame with PPS. The sorting induces an implicit stratification and set of joint probabilities of selection that can be effective in reducing sampling variance.

31,284 dwelling units in Wave 1. Because the segments selected for Waves 2 and 3 were listed in the fall of 1999, the special procedures used in Wave 1 were not employed for these waves.

A.1.1.2 Selection of the Permit Screening Sample

A separate building permit sample was drawn for each of the three initial waves of NSPY. The use of building permit samples was designed to offset the increase in variance that could potentially occur from the use of outdated measures of size in sample selection. Since the area segments were selected with probabilities proportionate to a size measure based on dwelling unit counts from the 1990 Census, the presence of new construction in the area sample can result in unduly high between-segment variances. To dampen the effect of new construction on sampling variances, only pre-1990 dwelling units were retained in the area sample. This was accomplished in the screening interview by asking the occupants when their dwelling unit was constructed and then terminating the screening process if the unit was built after April 1, 1990. To extend coverage to newer dwelling units not covered by the area sample, a separate sample of post-April 1990 dwelling units was drawn from a frame of building permits.

Sampling from building permits is feasible because (a) most localities require that a permit be obtained before building a residential structure and (b) the U.S. Census Bureau conducts a regular census of such building permit activity through its Census of Local Governments. Information on building permit activity is collected by the Census Bureau on a monthly basis for active permit-issuing offices and annually for less active offices. The census data thus can be used to construct a frame from which specific building permit offices can be drawn.

The stages of building permit sampling were similar to those used to select the area sample, except that there were five stages of sampling instead of three. First, only building permits issued within the 90 PSUs selected for the NSPY were eligible for sampling. Thus, a sample of building permit offices (BPOs) was initially selected within each of the sampled PSUs. At the next stage, a sample of "segments" was selected from each of the sampled BPOs, where here a segment is defined to be the set of permits issued by the office within a specific time interval. At the fourth stage, individual building permits were selected from the sampled segments. After selection of the building permits, a lister visited the building sites covered by the selected permits to list all of the dwelling units located at the site. After creating a list of dwelling units within each of the sampled segments, a sample of dwelling units was selected.

The number of dwelling units to be sampled from the building permit frame was set so that the proportion of the total sample selected through the permit frame would roughly equal the proportion of the national housing stock that was built between April 1, 1990, and the end of 1998. Statistics from the U.S. Census Bureau indicated that about 10 percent of the housing stock as of the end of 1998 met this criterion. The number of dwelling units sampled from the building permit frame for Wave 1 was 3,407, which was equal to about 10 percent of the total initial sample size. For Wave 2, there were 2,875 dwelling units drawn from the building permit frame, compared to 20,125 dwelling units from the area frame. Because the building permit frames constructed for all three recruitment waves covered dwelling units that were issued permits through the end of 1998, there was no coverage of dwelling units built in 1999 or 2000 for Wave 2. For Wave 3, there were 3,052 dwelling units drawn from the building permit frame, compared to 20,248 dwelling units from the area frame. For Wave 3, there was no coverage of new dwelling units that were permitted and built in 1999, 2000, or the first

half of 2001. Table A-1 summarizes the numbers of dwelling units selected for each of the three recruitment waves by source.

Table A-1. Number of dwelling units selected for the NSPY by wave and source

Source of sample	Wave 1	Wave 2	Wave 3
Area segments	30,993	20,125	20,248
Building permits	3,407	2,875	3,052
Missed DU Procedure	192	---	---
Missed mobile homes	99	---	---
Total	34,691	23,000	23,300

A.1.2 Household Screening and Subsampling (Waves 1 through 3)

Household screening and subsampling were used to identify eligible households and to oversample those with specific compositions to satisfy precision requirements for the three youth age ranges. In eligible households, one youth was selected from each age range represented, but no more than a total of two youth were selected per household. The parents and caregivers of the sampled youth were also identified and one was randomly selected for an extended interview. The practice of sampling up to two youth per household had the effect of concentrating the youth interviews in a smaller number of households than would be expected if sampling were conducted independently for each age range. This meant that youth in the less rare age domains were sampled at a higher rate if they happen to have a sibling in a rare age domain. This approach was particularly advantageous for NSPY because a sampled parent could report for more than one sampled child. Thus, concentrating the youth sample in a smaller set of households generated a more efficient parent sample.

To carry out the sampling efficiently, eligible households were assigned to three sampling strata based on the ages of the youth in the household. Because youth aged 12 to 13 were the rarest age domain, households containing such youth were always selected. They were thus placed into a stratum by themselves. Youth aged 9 to 11 were the next rarest domain. Households that contained a 9- to 11-year-old but no 12- or 13-year-olds were subsampled at Wave 1 and thus constituted a second stratum. Finally, 14- to 18-year-olds represented the most common age domain and were subsampled at the lowest rate. Households with at least one 14- to 18-year-old but no 9- to 13-year-olds constituted a third stratum. In summary, the following three sampling strata were used to select eligible households:

- Stratum A: Households containing at least one youth aged 12 to 13;
- Stratum B: Households containing at least one youth aged 9 to 11 but no youth aged 12 to 13; and
- Stratum C: Households containing at least one youth aged 14 to 18 but no youth aged 9 to 13.

Table A-2 summarizes estimates of the youth population by stratum for Wave 1. Tables A-3 and A-4 give the corresponding results for Waves 2 and 3, respectively. These estimates were derived using youth weights that had been poststratified (calibrated) to population estimates obtained from the Current Population Survey (e.g., see Section A.2.5 for a description of NSPY weighting procedures). The retention rates shown in the tables represent the percentages of the screened households of the given type to be retained for the sample. As can be seen in the tables, the retention rates originally specified in Wave 1 were subsequently modified for Waves 2 and 3.

Table A-2. Retention rates and estimated numbers of households and youth by household stratum: NSPY Wave 1

Household composition	Retention rate (%)	Households	Youth by age domain			
			9-11	12-13	14-18	Total 9-18
At least one 12- to 13-yr.-old	100	7,770,932	3,217,415	7,778,731	3,816,436	14,812,582
At least one 9- to 11-yr.-old but no 12- to 13-yr.-olds	70	8,449,930	9,309,863	0	3,075,451	12,385,315
At least one 14- to 18-yr.-old but no 9- to 13-yr.-olds	45	9,545,207	0	0	12,223,950	12,223,950
Total		25,766,069	12,527,278	7,778,731	19,115,837	39,421,846

Note: Estimates in this table are based on weights originally computed in Wave 1. Due to subsequent updates and revisions, the final weights used in analysis differ from those used to generate the estimates in this table.

Table A-3. Retention rates and estimated numbers of households and youth by household stratum: NSPY Wave 2

Household composition	Retention rate (%)	Households	Youth by age domain			
			9-11	12-13	14-18	Total 9-18
At least one 9- to 13-yr.-old	100	16,032,452	12,600,343	7,993,378	7,270,029	27,863,751
At least one 14- to 18-yr.-old but no 9- to 13-yr.-olds	45	9,344,405	0	0	12,067,622	12,067,622
Total		25,376,856	12,600,344	7,993,378	19,337,651	39,931,373

Note: Estimates in this table are based on weights originally computed in Wave 2. Due to subsequent updates and revisions, the final weights used in analysis differ from those used to generate the estimates in this table.

Table A-4. Retention rates and estimated numbers of households and youth by household stratum: NSPY Wave 3

Household composition	Retention rate (%)	Households	Youth by age domain			
			9-11	12-13	14-18	Total 9-18
At least one 9- to 13-yr.-old	100	16,163,113	12,825,995	8,055,046	8,425,940	29,306,981
At least one 14- to 18-yr.-old but no 9- to 13-yr.-olds	45	9,738,613	0	0	10,991,740	10,991,740
Total		25,901,726	12,825,995	8,055,046	19,417,680	40,298,721

Note: Estimates in this table are based on weights originally computed in Wave 3. Due to subsequent updates and revisions, the final weights used in analysis differ from those used to generate the estimates in this table.

Operationally, the subsampling of eligible households proceeded as follows. At the time the dwelling units were selected from the area and building permit segments, they were randomly assigned to one of three sampling rules:

Rule A. Interview if the household belongs to stratum A;

Rule AB. Interview if the household belongs to stratum A or B; or

Rule ABC. Interview if the household belongs to stratum A, B, or C.

If a dwelling unit was assigned sampling rule A, the interviewer was instructed to induct the household into the sample only if it contained a youth aged 12 or 13. If a dwelling unit was assigned sampling rule AB, the interviewer was instructed to induct the household into the sample if it contained one or more youth aged 9 to 13. If a dwelling unit was assigned sampling rule ABC, the interviewer was instructed to induct the household into the sample if there were any youth aged 9 to 18. The interviewer used a hard-copy screening questionnaire and simple focused questions to determine the presence of youth in the specified age ranges.

The proportions of dwelling units randomly assigned to the three sampling rules are shown in Tables A-5 through A-7 and were designed to achieve the retention rates indicated earlier in Tables A-2 through A-4. The corresponding eligibility rates shown in the tables were estimated from results obtained from the three initial recruitment waves. These rates are lower than those derived from the Current Population Survey (CPS). The differences are consistent with the roughly 30 percent undercoverage of eligible households achieved in all three recruitment waves (see Table A-21).

Table A-5. Wave 1 eligibility rates

Screener group	Screener sample (%)	Estimated age eligibility rate (%)	CPS predictions of eligibility rates (%)
A	30.1	5.6	7.5
AB	24.9	10.8	15.2
ABC	45.0	19.9	24.4
Total	100.0	12.2	17.0

Table A-6. Wave 2 eligibility rates

Screener group	Screener sample (%)	Estimated age eligibility rate (%)	CPS predictions of eligibility rates (%)
A-AB	55.1	10.9	15.7
ABC	44.9	17.0	24.9
Total	100.0	13.6	19.8

Table A-7. Wave 3 eligibility rates

Screener group	Screener sample (%)	Estimated age eligibility rate (%)	CPS predictions of eligibility rates (%)
A-AB	55.1	10.1	15.8
ABC	44.9	16.0	25.4
Total	100.0	13.3	20.1

For Waves 2 and 3, households in stratum B were sampled at the same rate as households in stratum A. The reason for this was to increase the sample size for youth aged 9 to 11. There was some concomitant increase in the sample size for youth aged 14 to 18. Operationally, this was accomplished by reassigning all households in screener group A to screener group AB. A larger sample size was desired for youth aged 9 to 11 at Waves 2 and 3 because of the decision to conduct followup interviews. Since there would be no new sample after Wave 3, the only way to achieve a sufficient sample of 12- to 13-year-olds after Wave 3 was to oversample the 9- to 11-year-olds at Waves 2 and 3.

Household screening was also used to account for multiple chances of selection of dwelling units built after the 1990 decennial census. As discussed earlier, most of these units had two chances of selection—once in the area sample and once in the building permit sample. This was true for all immobile units built after the 1990 Census in permit-issuing jurisdictions in Waves 2 and 3. For Wave 1, it was true only for immobile units built after the 1990 Census but before the original listing in late 1991. To determine these extra chances of selection, the screener included questions on the year the dwelling unit was built.

The only chance of selection for mobile homes was through the area sample because the building permit frame did not cover these types of dwellings. Therefore, all mobile homes in the area sample were retained for the study regardless of year built. However, for Wave 1, an additional procedure was needed because the area sample had not been updated since 1991. To extend coverage to new or missed mobile homes, interviewers recanvassed a subsample of the area segments to identify all currently existing mobile homes. Any mobile homes found in this process were compared with the old listing sheets to see whether they had been previously enumerated. All previously unlisted mobile homes were added to the sample. This procedure yielded a sample of 99 new or missed mobile homes.

Another activity that took place during the screening process for Wave 1 was called the missed dwelling unit (DU) procedure. At every single-family home, the interviewer asked whether there was a separate apartment in the basement, garage, or elsewhere. If such an apartment was found, the interviewer checked the original listing sheets to determine whether the apartment had been listed. If it had not been listed, the apartment was automatically added to the screening sample. A similar procedure was carried out in a sample of multifamily housing structures. If the first listed unit in the building was selected for the screening sample, the interviewer conducted a thorough recanvass of the structure to identify units missed by the original lister. Any previously unlisted apartments were added to the screening sample. This procedure generated a sample of 192 missed dwelling units for Wave 1.

The missed mobile home and missed DU procedures were not used for Waves 2 and 3. The listings used for those waves were compiled in mid-1999, making them fairly current for sampling in late 2000 and early 2001. Because of the screening and sampling procedures, all types of dwelling units built after 1998 were excluded from the sampling frames. In addition, all mobile homes put in place after the listing period in mid-1999 had no chance of selection for Wave 2 or 3.

A.1.3 Selection of Initial Samples for Waves 1 through 3

The procedure for Waves 1 through 3 was to prepare a list of eligible youth in each sampled household and sample one youth within each nonempty age range, subject to a maximum of two sample youth per household. In a household containing youth in all three of the age ranges, one youth from the 12-to-13 range was always selected. One of the remaining two age ranges was then chosen randomly, and a second youth was sampled from the selected age range. Within an age range, all

youth in the household had the same probability of selection. The interviewers then determined the relationship of all adults in the household to each sampled youth and also the relationship between sampled youth if two were selected. If two sampled youth were siblings, one adult was randomly selected from the set of adults in the household who were classified as the parents or caregivers of either youth. If two nonsiblings were selected, one adult was selected from each set of associated parents and caregivers. All of these procedures were accomplished with the aid of a CAPI instrument.

It should be noted that parents (caregivers) were randomly selected for the study rather than selecting the most knowledgeable or cooperative parent. Because the most knowledgeable and cooperative parent in two-parent households is often the mother, nonrandom selection would have resulted in a sample consisting mostly of mothers with very little data on fathers. To be able to measure the impact of the Campaign on fathers as well as mothers, random selection of parents was used.

For purposes of the study, parents included biological, adoptive, step, and foster parents living with the youth. Caregivers were defined to be persons serving in parental roles for youth who did not live with their natural parents. Some distinctions were made between these categories for sampling purposes. For example, stepparents were considered to be parents for sampling purposes only if they had lived with their stepchild for at least 6 months. Henceforth, in this discussion, the term parent will be used to refer to both parents and caregivers unless otherwise specified.

For sampling purposes, a youth who lived with both divorced or separated parents was assigned to the household where the youth spent the majority of the year. The only parents eligible for selection were the parents with whom the youth spent most of the year. It was possible to select a stepparent rather than a natural/adoptive parent of a sampled youth.

In the case of youth living with adults who were not their parents (under the strict definition of parents given above), special rules for sampling caregivers were implemented. For youth who were not emancipated[2] but lived with adults other than their parents, one or more primary caregivers in the household were identified. These caregivers may or may not have been the youth's legal guardians.[3] If there were more than one resident primary caregiver, one was randomly selected for the parent interview.

For emancipated youth living separately from their parents, a caregiver was generally not required. However, when there was an adult present who might be a caregiver (such as a grandmother), it was determined whether that adult was a caregiver and, if so, an attempt was made to recruit him or her for a parent interview.

Youth under age 19 who were serving in parental roles (e.g., an older sibling in a pair of orphans or a teenage stepmother) were considered ineligible for the youth selection but eligible for the parent selection.

As mentioned previously, youth residing in group quarters were not eligible for selection in any of the three recruitment waves. Thus, youth living in boarding schools and college dormitories were excluded from the scope of the survey. This exclusion was made because it was felt that dormitory residents could not be easily interviewed at their parents' homes and that their experiences were so

[2] The criteria for identifying emancipated youth vary by state but generally involve age and marital status.

[3] If the caregiver was not the legal guardian, a parent interview was conducted with the caregiver and the legal guardian was contacted for permission to interview the youth.

different from the majority of youth that they would have to be analyzed separately. During screening, the interviewer specifically asked respondents not to count these youth as household members. Despite the exclusion of dormitory residents, college students under 19 years of age who were living at home or in private apartments were eligible for sampling. It was decided that a broader exclusion of college students was not necessary for analytic purposes and would render the remaining sample of 18-year-olds unrepresentative of the universe that most data users would expect to find. The exclusion of dormitory residents did pose some special challenges in the weighting process described in Section A.2.5.

Table A-8 summarizes counts of interviewed youth at Wave 1 by age and by household stratum. Tables A-9 and A-10 give the corresponding counts for Waves 2 and 3, respectively. Within households completing the household roster, person-level (conditional) response rates were generally high. Section B.8 of Appendix B provides additional details about the response rates achieved for Waves 1 through 3.

Table A-8. Rostered households and completed parent and youth interviews by household stratum for NSPY Wave 1

			Youth age domain			
Household composition	Rostered households	Parents	9-11	12-13	14-18	Total
At least one 12- to 13-yr.-old	1,183	1,053	344	1,053	412	1,809
At least one 9- to 11-yr.-old but no 12- to 13-yr.-olds	826	732	742	0	228	970
At least one 14- to 18-yr.-old but no 9- to 13-yr.-olds	593	499	0	0	519	519
Total	2,602	2,284	1,086	1,053	1,159	3,298

Note: All counts reflect results of final data cleaning activities and may differ slightly from those in previously published reports.

Table A-9. Rostered households and completed parent and youth interviews by household stratum for NSPY Wave 2

			Youth age domain			
Household composition	Rostered households	Parents	9-11	12-13	14-18	Total
At least one 9- to 13-yr.-old	1,498	1,329	922	658	448	2,028
At least one 14- to 18-yr.-old but no 9- to 13-yr.-olds	368	303	0	0	333	333
Total	1,866	1,632	922	658	781	2,361

Note: All counts reflect results of final data cleaning activities and may differ slightly from those in previously published reports.

Table A-10. Rostered households and completed parent and youth interviews by household stratum for NSPY Wave 3

			Youth age domain			
Household composition	Rostered households	Parents	9-11	12-13	14-18	Total
At least one 9- to 13-yr.-old	1,607	1,424	977	727	475	2,179
At least one 14- to 18-yr.-old but no 9- to 13-yr.-olds	322	258	0	0	279	279
Total	1,929	1,682	977	727	754	2,458

Note: All counts reflect results of final data cleaning activities and may differ slightly from those in previously published reports.

A.1.4 Selection of Followup Samples

Wave 4

Under the NSPY sample design, subsamples of youth and parents selected for the initial recruitment waves (i.e., Waves 1 through 3) were retained for followup in subsequent data collection waves. No new samples were selected for any of the followup waves. For Wave 4, the first followup of Wave 1, all youth and parents in households that completed the screener roster in Wave 1 were included in the followup sample if the household contained at least one Wave 1 respondent (either youth or parent). Note that under the selection criterion employed for Wave 4, a small number of youth and parents who were selected for Wave 1 but who did not complete a Wave 1 interview were refielded in Wave 4. The "extra" youth and parents who were obtained in Wave 4 were used only for cross-sectional analyses at Wave 4. Table A-11 summarizes counts of youth and parents completing the Wave 4 interview by age of sampled youth. Details about the response rates achieved in Wave 4 are given in Section B.8 of Appendix B.

Table A-11. Number of completed parent and youth interviews by youth age domain and wave

		Youth age domain				
Wave	Parents	9-11	12-13	14-18	12-18	Total
1	2,230	1,086	1,055	1,157	2,212	3,298
2	1,617	922	658	781	1,439	2,361
3	1,657	977	727	754	1,481	2,458
4	1,743	438	663	1,376	2,039	2,477
5	2,872	988	1,209	1,842	3,051	4,039
6	1,636	127	618	1,522	2,140	2,267
7	2,616	384	1,124	2,079	3,203	3,587
8	1,483	0	413	1,570	1,983	1,983
9	2,377	12	880	2,251	3,131	3,143
Total	18,231	4,934	7,347	13,332	20,679	25,613

Note: All counts reflect results of final data cleaning activities and may differ slightly from those in previously published reports.

Wave 5

For Wave 5, the first followup of Waves 2 and 3, all youth and parents in households that completed the screener roster in Waves 2 and 3 were included in the followup sample if the household contained at least one respondent from the prior wave (either youth or parent). Under this selection criterion, a small number of youth and parents who were selected but did not complete a Wave 2 or 3 interview were refielded in Wave 5. The "extra" youth and parents who were obtained in Wave 5 were used only for cross-sectional analyses at Wave 5. Table A-11 summarizes counts of youth and parents completing the Wave 5 interview by age of sampled youth. Details about the response rates achieved in Wave 5 are given in Section B.8 of Appendix B.

Wave 6

For Wave 6, the second followup of Wave 1, only those youth and parents who were expected to be eligible for NSPY at Wave 6 *and* who met the following criteria were refielded: (a) the youth/parent resided in a household in which at least one sampled youth completed either the Wave 1 or Wave 4 interview, and (b) the household was neither "not locatable" nor "out of area" in Wave 4. Under these selection criteria, a small number of youth and parents who did not complete both the Wave 1

and Wave 4 interviews were refielded in Wave 6. The "extra" youth and parents who were obtained in Wave 6 were used only for cross-sectional analyses at Wave 6. Table A-11 summarizes counts of youth and parents completing the Wave 6 interview by age of sampled youth. Details about the response rates achieved in Wave 6 are given in Section B.8 of Appendix B.

Wave 7

For Wave 7, the second followup of combined Waves 2 and 3, only those youth and parents who were expected to be eligible for NSPY at Wave 7 *and* who met the following criteria were refielded: (a) the youth/parent resided in a household in which at least one sampled youth completed either the Wave 2/3 or Wave 5 interview, and (b) the household was neither "not locatable" nor "out of area" in Wave 5. Under these selection criteria, a small number of youth and parents who did not complete both the Wave 2/3 and Wave 5 interviews were refielded in Wave 7. The "extra" youth and parents who were obtained in Wave 7 were used only for cross-sectional analyses at Wave 7. Table A-11 summarizes counts of youth and parents completing the Wave 7 interview by age of sampled youth. Details about the response rates achieved in Wave 7 are given in Section B.8 of Appendix B.

Wave 8

For Wave 8, the third followup of Wave 1, only those youth and parents who were expected to be eligible for NSPY at Wave 8 *and* who met the following criteria were refielded: (a) the youth/parent resided in a household in which at least one sampled youth completed either the Wave 1, 4, or 6 interview, and (b) the household was neither "not locatable" nor "out of area" in Wave 6. Under these selection criteria, a small number of youth and parents who did not complete all three of the prior interviews were refielded in Wave 8. The "extra" youth and parents who were obtained in Wave 8 were used only for cross-sectional analyses at Wave 8. Table A-11 summarizes counts of youth and parents completing the Wave 8 interview by age of sampled youth. Details about the response rates achieved in Wave 8 are given in Section B.8 of Appendix B.

Wave 9

For Wave 9, the third followup of Waves 2 and 3 combined, only those youth and parents who were expected to be eligible for NSPY at Wave 9 *and* who met the following criteria were refielded: (a) the youth/parent resided in a household in which at least one sampled youth completed either the Wave 2/3, 5, or 7 interview, and (b) the household was neither "not locatable" nor "out of area" in Wave 7. Under these selection criteria, a small number of youth and parents who did not complete all three of the prior interviews were refielded in Wave 9. The "extra" youth and parents who were obtained in Wave 9 were used only for cross-sectional analyses at Wave 9. Table A-11 summarizes counts of youth and parents completing the Wave 9 interview by age of sampled youth. Details about the response rates achieved in Wave 9 are given in Section B.8 of Appendix B.

A.2 Development of Weights

An analysis weight was calculated for each completed interview. Different weights were prepared for different types of analyses. For Waves 1 through 9, separate sets of national cross-sectional weights were developed for youth and youth-parent dyads. For Waves 4 through 9, separate longitudinal weights were created in addition to the cross-sectional weights. All of the weights were designed to reflect overall selection probabilities and to compensate for nonresponse and undercoverage. The adjustments for undercoverage involved a post-stratification process called ratio raking (or simply

"raking") in which the nonresponse-adjusted weights were calibrated in such a manner that the resulting weighted counts for selected domains agreed with the corresponding "known" population counts from independent sources. The final weight for a respondent, after nonresponse and raking adjustments, can be viewed as the number of population members that each respondent represents. Details about the weighting process are given in the following sections.

A.2.1 Dwelling Unit Base Weights

Base weights are used to reflect a person's probability of selection into the sample. The base weight is defined to be the reciprocal of the probability of selection. Thus, persons with small probabilities of selection have large base weights and those with large probabilities have small base weights. If there were no nonresponse or undercoverage, these base weights would yield unbiased estimates of population parameters such as the percent of youth who engage in a particular behavior.

The base weights reflected the probability of selection at each stage of sampling: PSU, segment, dwelling unit (DU), and person. The calculation of these probabilities at each stage was straightforward. However, since the selection of persons could be carried out only in households where the screener was completed, the person-level base weight also included an adjustment for household nonresponse.

For each of the three recruitment waves, Waves 1 through 3, the base weight for a dwelling unit (DU) was generally computed as:

$$BW_{DU} = \frac{1}{\Pr\{\text{PSU}\}\ \Pr\{\text{segment} \mid \text{PSU}\}\ \Pr\{\text{DU} \mid \text{segment}\}} \qquad (1)$$

where $\Pr\{\text{PSU}\}$ is the probability of selecting the PSU, $\Pr\{\text{segment}|\text{PSU}\}$ is the probability of selecting the segment within the PSU, and $\Pr\{\text{DU}|\text{segment}\}$ is the probability of selecting the dwelling unit within the segment.

For building permit segments (see Section A.1.1.2), the dwelling unit base weight defined above also included adjustments for undercoverage due to the failure to locate permits for some segments and the lack of coverage of new housing units in jurisdictions where building permits were not required. The required adjustments, which varied by Census region, were derived from statistics in the Census Bureau's reports on construction starts.

A.2.2 Adjustment of Dwelling Unit Weights for Screener Nonresponse

Next, the dwelling unit base weights were adjusted for screener nonresponse in two phases. The first phase, referred to as the "doorstep nonresponse adjustment," was designed to compensate for those DUs for which it was not possible to determine whether eligible youth were present at the sampled address. For this adjustment, a dwelling unit was considered to be a respondent if information about the presence or absence of eligible children had been collected from either the occupants of the household or from their neighbors. The second phase adjustment, referred to as the "roster nonresponse adjustment," was designed to compensate for households known to contain eligible youth, but for which it was not possible to prepare a household roster and select a sample of youth (and parents). For this adjustment, a DU was considered to be a respondent if it contained any eligible children *and* an adult resident who was willing to provide a roster of the occupants of the household,

their ages, and their relationships to the household members. Otherwise, the DU was classified as a nonrespondent.

The dwelling unit base weights were adjusted to compensate for unequal response rates across different groups in the sample. Special data mining software referred to as "multiple additive regression trees" (MART) was used to form the groups required for the nonresponse adjustments (see Friedman, 1999, for a description of MART). The MART algorithm, which uses predictive data mining techniques, was applied to identify subgroups that are homogeneous with respect to response propensity. About 60 block-group-level variables describing characteristics of the neighborhood taken from the 2000 Population Census were specified as predictor variables in the MART software.[4] Some examples of the variables used as predictors include percent of persons in various age groups, percent of persons of certain race and ethnic groups, percent of households that own (rather than rent) their homes, percent of persons residing in mobile homes, percent of persons who are U.S. citizens, and percent of households with incomes below the poverty level. Within the groups (referred to as "weighting classes") formed by the MART algorithm, the corresponding first-phase (doorstep) and second-phase (roster) weighted response rates were computed and used to obtain the nonresponse-adjusted dwelling unit weight. That is, the first-phase nonresponse-adjusted weight for the *i*th dwelling unit in first-phase weighting class *h* was computed as:

$$BW_{hi}^{Adj1} = F_h^{(1)} \, BW_{DU,i} \tag{2}$$

where $BW_{DU,i}$ is the base weight for dwelling unit *i* and $F_h^{(1)}$ is the reciprocal of the first-phase (doorstep) weighted response rate in class *h*. Next, using the adjusted weights calculated above, the final nonresponse-adjusted weight for the *i*th dwelling unit in second-phase weighting class *k* was computed as:

$$BW_{ki}^{Adj} = F_k^{(2)} \, BW_{ki}^{Adj1} \tag{3}$$

where BW_{ki}^{Adj1} is the first-phase nonresponse-adjusted weight for dwelling unit *i* in second-phase weighting class *k* and $F_k^{(2)}$ is the reciprocal of the second-phase (roster) weighted response rate in class *k*.

After adjustment for both phases of screener nonresponse, the weights were further adjusted for screener-based subsampling. As indicated in Section A.1.2, dwelling units were randomly retained for the study at rates that depended on household composition. For example, dwelling units with at least one youth 12 to 13 years old were retained with certainty. On the other hand, dwelling units with youth 14 to 18 years old but no youth 9 to 13 years old were retained at a rate of 45 percent (e.g., see Tables A-2 through A-4). Thus, in general, the final weight for DU *i* in stratum *c* was computed as:

$$BW_{ci}^{Fin} = BW_{ci}^{Adj} \; / \; R_c \tag{4}$$

where BW_{ci}^{Adj} is the nonresponse-adjusted weight for DU *i* in stratum *c*, and R_c is the corresponding retention rate.

[4] In previous reports, the weights used in analysis were derived using block group-level variables based on 1990 Census data, the most current data available at the time, With the availability of the 2000 Census data, all of the 1990 Census-based block group-level variables used in weighting adjustments were replaced by their 2000 Census versions. This change resulted in somewhat better statistical adjustments.

Table A-12 summarizes results of the doorstep (first phase) nonresponse adjustment process by recruitment wave. The weighted counts shown in the first row of the table are estimates of the number of occupied dwelling units based on the sampled DUs. The (weighted) response rates shown in the table represent the proportion of occupied DUs for which eligibility for the survey was determined. The important predictors of response propensity included URBAN_C (level of urbanization of the sampled area), MULTIDU (whether the DU was single- or multifamily), and other block group level variables derived from 2000 Census data. The number of weighting cells formed for doorstep nonresponse adjustment and the range of weighting adjustment factors were generally similar for all three waves.

Table A-12. Summary of doorstep nonresponse adjustment process for Wave 1 through 3.

	Recruitment waves (Round 1)		
Doorstep NR adjustment (1st Phase)	Wave 1	Wave 2	Wave 3
Base-weighted occupied DU count*	99,570,000	109,407,000	105,849,000
Weighted count of DUs for which eligibility was determined*	94,735,000	104,740,000	101,360,000
Weighted response rate*	95%	96%	96%
Top 5 MART predictors†	URBAN_C	URBAN_C	URBAN_C
	MULTIDU	PCT3_52	MULTIDU
	PCT3_23	PCT3_49	PCT3_22
	PCT3_11	PCT3_8	PCT3_57
	PCT3_19	PCT3_31	PCT3_37
Number of NR (MART) cells formed	13	12	11
Range of adjustment factors	1.01 to 1.56	1.01 to 1.72	1.01 to 1.66

* Weights are the dwelling unit base weights defined by formula (1).

† URBAN_C indicates the type of area: 0 = rural, 1 = town or exurban area, 2 = second city of metropolitan area, 3 = suburb of metropolitan area, 4 = urban part of metropolitan area.

MULTIDU indicates the multifamily status of dwelling unit: 0 = single family, 1 = first unit of a multifamily dwelling, 2 = all other units of a multifamily dwelling, blank = not available. For wave 1, a blank indicates that DU is part of the building permit sample.

PCT3_52 = tercile rank of the segment with respect to average number of persons per room; PCT3_23 = tercile rank of the segment with respect to percent of nonfamily households; PCT3_49 = tercile rank of the segment with respect to percent of housing units that are vacant for occasional use; PCT3_22 = tercile rank of the segment with respect to percent of households with children under 18 headed by female; PCT3_11 = tercile rank of the segment with respect to percent of persons who are white; PCT3_8 = tercile rank of the segment with respect to percent of persons who live outside of urbanized areas; PCT3_57 = tercile rank of the segment with respect to average number of persons per vehicle; PCT3_19 = tercile rank of the segment with respect to percent of persons who are Cuban American; PCT3_31 = tercile rank of the segment with respect to percent of foreign born noncitizens; and PCT3_37 = tercile rank of the segment with respect to percent of persons in labor force who are employed.

Table A-13 summarizes the corresponding results of the roster (second phase) nonresponse adjustment process by recruitment wave. The weighted counts shown in this table are estimates of the number of occupied dwelling units with youth 9 to 18 years of age. The (weighted) response rates shown in the table represent the proportion of occupied DUs with youth 9 to 18 years of age for which a completed household roster was obtained. Again, the important predictors of response propensity included URBAN_C (level of urbanization of the sampled area) and MULTIDU (whether the DU was single- or multifamily). The number of weighting cells formed for doorstep nonresponse adjustment and the range of weighting adjustment factors were generally similar for all three waves.

Table A-13. Summary of roster nonresponse adjustment process for Waves 1 through 3.

Roster NR adjustment (2nd Phase)	Recruitment waves (Round 1)		
	Wave 1	Wave 2	Wave 3
NR-adjusted wtd. count of eligible DUs*	17,168,622	17,149,583	16,242,865
NR-adj. wtd. count of DUs providing roster*	12,876,750	12,862,500	12,182,250
Weighted response rate*	75%	75%	75%
Top 5 MART predictors†	URBAN_C	URBAN_C	URBAN_C
	MULTIDU	PCT3_48	PCT3_58
	PCT3_39	PCT3_36	PCT3_34
	PCT3_13	PCT3_39	PCT3_23
	PCT3_1	MULTIDU	PCT3_55
Number of NR (MART) cells formed	5	4	5
Range of adjustment	1.10 to 1.60	1.09 to 1.63	1.14 to 1.62

* Weights are the first-phase nonresponse-adjusted weights defined by formula (2).

† URBAN_C indicates the type of area: 0 = rural, 1 = town or exurban area, 2 = second city of metropolitan area, 3 = suburb of metropolitan area, 4 = urban part of metropolitan area.

MULTIDU indicates the multifamily status of dwelling unit: 0 = single family, 1 = first unit of a multifamily dwelling, 2 = all other units of a multifamily dwelling, blank = not available. For Wave 1, a blank indicates that DU is part of the building permit sample.

PCT3_1 = tercile rank of the segment with respect to percent of persons 9-18 years of age; PCT3_13 = tercile rank of the segment with respect to percent of persons who are native American; PCT3_23 = tercile rank of the segment with respect to percent of nonfamily households; PCT3_34 = tercile rank of the segment with respect to percent of persons with bachelors degree or higher in segment; PCT3_36 = tercile rank of the segment with respect to percent of persons 16+ who are employed; PCT3_39 = tercile rank of the segment with respect to percent of persons working in manufacturing; PCT3_48 = tercile rank of the segment with respect to percent of vacant dwelling units; PCT3_55 = tercile rank of the segment with respect to percent of single family structures; and PCT3_58 = tercile rank of the segment with respect to percent of renters with no cash rent.

A.2.3 Initial Youth Weights

As mentioned in Section A.1.3, there were three age classes of special analytic interest for youth sampling purposes: 9 to 11, 12 to 13, and 14 to 18. If there were youth present in all three age classes, the first step in the youth subsampling was to select two out of the three age ranges. The 12-to-13 age range was selected with certainty. One of the remaining two age ranges was selected with equal probability. Thus, the first component in the youth probability of selection for youth aged 9 to 11 or 14 to 18 in households with all three age classes was a factor of 0.5. Next, one youth was selected from each of the selected age ranges. For example, if there were four youth present in the given age range, the probability of selection within the range was 1/4 = 0.25. The two factors were multiplied together to create the probability of selecting the *j*th youth in dwelling unit *i*:

$$Pr\{\text{youth } j \mid \text{DU } i\}. \tag{5}$$

The initial youth weight for the *j*th youth in dwelling unit *i* was then calculated as the final dwelling unit weight divided by the within-dwelling unit probability of selecting the youth, i.e.,

$$IW_{ij}^{youth} = BW_i^{Fin} \; / \; \Pr\{\text{youth } j \mid \text{DU } i\}. \tag{6}$$

Because no new youth were selected in Round 2 (Waves 4 and 5), Round 3 (Waves 6 and 7), or Round 4 (Waves 8 and 9), the initial weights calculated for the recruitment waves were also the initial weights for youth in the followup rounds. As described below, the wave-specific initial youth weights

were adjusted for nonresponse and then ratio-adjusted to independent population totals corresponding to each wave.

A.2.4 Nonresponse Adjustment of Youth Weights

Responding youth were defined to be those who answered the last question on general ad exposure (D13 in the youth questionnaire). Nonresponding youth were those who broke off the self-administered data collection prior to this point, or else failed to start the data collection. Nonrespondents included sampled youth whose parents refused consent or otherwise failed to provide consent, those who refused personal assent, and those who were never reached to do the interview for any reason. Among those who did not complete the questionnaire, a difference was drawn between those who were physically or mentally incapable of completing the interview and those who simply chose not to. The first group was considered to be ineligible sample youth rather than nonresponding sample youth. The distinction matters because the weight of ineligible youth are not redistributed to responding youth through the nonresponse adjustments described below. Youth who could not communicate in English or Spanish were also considered to be ineligible for the evaluation since the Campaign's television and radio ads were broadcast only in these languages. Also included in the ineligible youth category were persons under 19 years of age who held parental roles for other youth aged 9 to 18. This might occur by reason of marrying an older person whose children are in the age range or by reason of caring for younger siblings. (Such persons were, however, eligible for the parent interview.)

The set of 60 block group-level variables used for doorstep and roster nonresponse adjustment plus selected youth and parent characteristics (i.e., age, gender, and race/ethnicity) and characteristics on household composition were used in conjunction with the MART software to develop a set of weighting cells for youth nonresponse adjustment. The household-level characteristics included items such as whether both parents lived in the household, whether the youth was an only child, the number of youth living in the household, and whether there was an unrelated person living in the household. All of these variables were obtained from the household roster. The set of nonresponse adjustment cells determined by the MART algorithm was then used to adjust the weights of the responding youth. That is, the nonresponse-adjusted weight for responding youth j in weighting class g was computed as:

$$W_{gj}^{NR} = G_g \, IW_{gj}^{youth} , \tag{7}$$

where IW_{gj}^{youth} is the initial youth weight and G_g is the reciprocal of the weighted response rate in weighting class g.

Tables A-14 through A-16 summarize results of the youth nonresponse adjustment process by wave. The weights used to compute the weighted counts given in the first two rows of each table are the initial youth weights described in Section A.2.3. The corresponding weighted response rates are defined to be the ratios of the weighted count of respondents to the weighted count of sampled youth. Since the denominator of the weighted response rate includes youth who were sampled from households completing the household enumeration (roster), the response rates are conditional rather than overall rates. Note that for Waves 4 through 9, the "sampled" youth are those that were refielded for followup according to criteria developed for each wave (e.g., see Section A.1.4).

Table A-14. Summary of youth nonresponse adjustment process for recruitment Waves 1 through 3

Youth[1] nonresponse adjustment	Recruitment waves (Round 1)		
	Wave 1	Wave 2	Wave 3
Weighted count of sampled youth[2]	27,529,801	27,343,697	26,084,536
Weighted count of responding youth[2]	24,801,496	25,111,988	23,740,560
Weighted (conditional) response rate[2,3]	90%	92%	91%
Top 5 MART predictors[4]	URBAN_C	URBAN_C	URBAN_C
	MULTIDU	NUMYUTH	BOTHPAR
	NUMYUTH	MULTIDU	MULTIDU
	PCT3_12	NONREL	PCT3_37
	PCT3_58	BOTHPAR	PCT3_1
Number of NR (MART) cells formed	5	8	9
Range of adjustment	1.02 to 1.54	1.01 to 1.60	1.02 to 1.48

[1] Youth 9 to 8 years of age.

[2] Weights are the initial youth weights defined by formula (6).

[3] Conditional on youth sampled from completed household rosters.

[4] URBAN_C indicates the type of area: 0 = rural, 1 = town or exurban area, 2 = second city of metropolitan area, 3 = suburb of metropolitan area, 4 = urban part of metropolitan area.

MULTIDU indicates the multifamily status of dwelling unit: 0 = single family, 1 = first unit of a multifamily dwelling, 2 = all other units of a multifamily dwelling, blank = not available. For Wave 1, a blank indicates that DU is part of the building permit sample.

NUMYUTH indicates the number of youth 9-18 years of age living in the household.

NONREL indicates whether or not an unrelated person lives in the household.

BOTHPAR indicates whether or not both parents live in the household.

PCT3_1 = tercile rank of the segment with respect to percent of persons 9-18 years of age; PCT3_12 = tercile rank of the segment with respect to percent of persons who are black; PCT3_37 = tercile rank of the segment with respect to percent of persons in labor force who are employed; and PCT3_58 = tercile rank of the segment with respect to percent of renters with no cash rent.

Table A-15. Summary of youth nonresponse adjustment process for followup Waves 4 through 7

	First followup (Round 2)		Second followup (Round 3)	
Youth[1] nonresponse adjustment	Wave 4	Wave 5	Wave 6	Wave 7
Wtd. count of sampled youth[2]	18,980,623	22,385,082	17,840,004	20,281,943
Wtd. count of responding youth[2]	17,635,056	19,692,896	15,668,010	17,593,974
Wtd. (cond.) response rate[2,3]	93%	88%	88%	87%
Top 5 MART predictors[4]	URBAN_C	YUTHRAC1	URBAN_C	INT_AGEP
	W4IMPAGE	PARRAC1	RACEETHP	URBAN_C
	PARRAC1	URBAN_C	AGEP_3	RACEETHP
	PCT3_37	MULTIDU	INT_AGEC	INT_AGEC
	NONREL	W5IMPAGE	PCT3_49	RACEETHC
Number of NR (MART) cells	13	9	7	8
Range of adjustment	1.00 to 1.43	1.05 to 1.35	1.02 to 1.28	1.03 to 1.50

[1] Youth 12 to 8 years of age.

[2] Weights are the initial youth weights defined by formula (6). Only those originally-sampled youth in households for which eligibility was ascertained in the given followup wave are included in the weighted counts. The percentage of households for which eligibility was ascertained in Round 2 was generally lower for Wave 4 than Wave 5.

[3] Conditional on youth refielded for followup.

[4] URBAN_C indicates the type of area: 0 = rural, 1 = town or exurban area, 2 = second city of metropolitan area, 3 = suburb of metropolitan area, 4 = urban part of metropolitan area.

MULTIDU indicates the multifamily status of dwelling unit: 0 = single family, 1 = first unit of a multifamily dwelling, 2 = all other units of a multifamily dwelling, blank = not available. For Wave 1, a blank indicates that DU is part of the building permit sample.

NONREL indicates whether or not an unrelated person also loves in the household.

BOTHPAR indicates whether or not both parents live in the household.

W4IMPAGE, W5IMPAGE, and INTAGEC indicate the age of youth at followup.

RACEETHC indicates race/ethnicity of youth.

PARRAC1 and RACEETHP indicate the race/ethnicity of parent.

AGEP_3 and INT_AGEP indicate the age of parent.

PCT3_37 = tercile rank of the segment with respect to percent of persons in labor force who are employed; and PCT3_49 = tercile rank of the segment with respect to percent of vacant housing units for occasional use.

Table A-16. Summary of youth nonresponse adjustment process for followup Waves 8 and 9

Youth[1] nonresponse adjustment	Third followup (Round 4)	
	Wave 8	Wave 9
Weighted count of sampled youth[2]	14,372,201	15,201,637
Weighted count of responding youth[2]	12,967,200	13,616,143
Weighted (conditional) response rate[2,3]	90%	90%
Top 5 MART predictors[4]	URBAN_C	URBAN_C
	INT_AGEP	RACEETHP
	MULTIDU	INT_AGEC
	RACEETHP	RACEETHC
	GENDERP	AGEP_3
Number of NR (MART) cells formed	6	7
Range of adjustment	1.04 to 1.50	1.03 to 1.55

[1] Youth 12½ to 8 years of age.

[2] Weights are the initial youth weights defined by formula (6).

[3] Conditional on youth refielded for followup.

[4] URBAN_C indicates the type of area: 0 = rural, 1 = town or exurban area, 2 = second city of metropolitan area, 3 = suburb of metropolitan area, 4 = urban part of metropolitan area.

MULTIDU indicates the multifamily status of dwelling unit: 0 = single family, 1 = first unit of a multifamily dwelling, 2 = all other units of a multifamily dwelling, blank = not available. For Wave 1, a blank indicates that DU is part of the building permit sample.

INT_AGEC indicates age of youth at followup.

RACEETHC indicates race/ethnicity of youth.

AGEP_3 and INT_AGEP indicate the age of parent.

RACEETHP indicates the race/ethnicity of parent.

GENDERP indicates sex of parent.

A.2.5 Poststratification (Raking) Adjustments

Raking is a form of poststratification adjustment that is commonly used to calibrate survey estimates to known population totals. The goals of raking are to reduce biases due to undercoverage and nonresponse and to reduce sampling errors. Raking may be thought of as an iterative form of poststratification in which the weights are successively ratio-adjusted to multiple sets of marginal control totals (referred to as "raking dimensions") until the resulting weighted sums equal the control totals specified for each dimension. The sample sizes associated with the levels of each raking dimension are the important determinants of the stability of the raking procedure, not the cells formed by a complete cross-classification of the variables defining the raking dimensions. This permits the use of more auxiliary variables than would be feasible with traditional or "direct" poststratification. For this reason, raking was used to calibrate the youth weights rather than direct poststratification.

The raking dimensions defined for the poststratification adjustments varied by wave to accommodate the aging of the NSPY sample. For Waves 1 to 3, the raking dimensions were:

- Gender crossed by three age groups (ages 9 to 11, 12 to 13, and 14 to 18); and

- Race/Ethnicity (non-Hispanic-non-Black, non-Hispanic-Black, Hispanic) crossed by three age groups (ages 9 to 11, 12 to 13, and 14 to 18).

For Waves 4 to 7, the raking dimensions were:

- Gender crossed by two age groups (ages 12 to 13, and 14 to 18); and
- Race/Ethnicity (non-Hispanic-non-Black, non-Hispanic-Black, Hispanic) crossed by two age groups (ages 12 to 13, and 14 to 18).

For Waves 8 and 9, the raking dimensions were:

- Gender crossed by "single" years of age (12½, 13, 14, ..., 18); and
- Race/Ethnicity (non-Hispanic–non-Black, non-Hispanic–Black, Hispanic) crossed by single years of age (12½, 13, 14, ..., 18).

Independent estimates of the total 9- to 18-year-old civilian population by gender, age group, and race/ethnicity were obtained from the Current Population Survey (CPS) for the relevant time periods covered by each wave. These population totals were then adjusted downward to exclude the civilian noninstitutional group quarters population, using estimates derived from the 1990 Census Public Use Microdata System (PUMS) files. The resulting adjusted population totals were used as control totals in the raking process. Note that the youth samples for Waves 8 and 9 were poststratified to the 12½- to 18-year old population because there were virtually no youth under 12½ years old by Waves 8 and 9.

After the youth nonresponse-adjusted weights were calculated as described earlier, the final ("raked") youth weights were computed as:

$$W_{kj}^{Fin} = H_k W_{kj}^{NR} \tag{8}$$

where W_{kj}^{NR} is the nonresponse-adjusted youth weight, and H_k is the final adjustment determined by the iterative raking process. The final weighted counts of responding youth are summarized in Tables A-17 through A-20 by wave. The estimated coverage rate, which is defined to be the ratio of the sum of the weights before raking to the corresponding (CPS-based) control total, is shown in Table A-21 by age, race, and gender for each of the three recruitment waves. Coverage rates were not computed for Waves 4 through 9, because the followup samples were subsets of the initial samples.

Table A-17. Unweighted and weighted counts of responding youth in recruitment waves for selected subgroups

	Recruitment waves (Round 1)					
	Wave 1		Wave 2		Wave 3	
Subgroup	Unwtd.	Wtd.	Unwtd.	Wtd.	Unwtd.	Wtd.
Total 9-18 years*	3,298	39,661,844	2,361	39,956,092	2,458	40,307,109
9 to 11	1,088	12,502,892	922	12,620,494	977	12,835,759
12 to 13	1,050	7,878,424	658	7,997,114	725	8,054,534
14 to 18	1,160	19,280,528	781	19,338,484	756	19,416,816
Male	1,729	20,198,810	1,200	20,422,666	1,257	20,677,608
Female	1,569	19,463,033	1,161	19,533,426	1,201	19,629,501
Non-Black, non-Hispanic	2,312	27,707,919	1,653	27,675,377	1,724	27,942,746
Black, non-Hispanic	478	6,164,046	371	6,294,252	363	6,266,066
Hispanic	508	5,789,879	337	5,986,463	371	6,098,297

* Note: Counts include all youth 9 to 18 years old. These are the ages included in the first through third semiannual reports. However, only those youth 12½ to 18 years of age are included in the analyses described in the present report. Also, the numbers in this table may differ slightly from those given in the main report due to corrections made after the weighting process was completed.

Table A-18. Unweighted and weighted counts of responding youth in first followup waves for selected subgroups

	First followup (Round 2)			
	Wave 4		Wave 5	
Subgroup	Unwtd.	Wtd.	Unwtd.	Wtd.
Total 12-18 years*	2,054	27,711,435	3,074	27,956,282
12 to 13	663	8,197,998	1,211	8,254,816
14 to 18	1,391	19,513,437	1,863	19,701,466
Male	1,091	14,186,399	1,542	14,303,437
Female	963	13,525,035	1,532	13,652,845
Non-Black, non-Hispanic	1,473	19,284,631	2,183	19,473,896
Black, non-Hispanic	269	4,364,886	454	4,363,462
Hispanic	312	4,061,917	437	4,118,924

* Note: Counts include all youth 12 to 18 years old. These are the ages included in the fourth and fifth semiannual reports. However, only those youth 12½ to 18 years of age are included in the analyses described in the present report. Also, the numbers in this table may differ slightly from those given in the main report due to corrections made after the weighting process was completed.

Table A-19. Unweighted and weighted counts of responding youth in second followup waves for selected subgroups

	Second followup (Round 3)			
	Wave 6		Wave 7	
Subgroup	Unwtd.	Wtd.	Unwtd.	Wtd.
Total 12-18 years*	2,140	28,272,984	3,203	28,600,478
12 to 13	619	8,465,097	1,123	8,464,347
14 to 18	1,521	19,807,887	2,080	20,136,131
Male	1,133	14,465,684	1,618	14,624,198
Female	1,007	13,807,300	1,585	13,976,280
Non-Black, non-Hispanic	1,537	19,451,884	2,278	19,640,478
Black, non-Hispanic	274	4,504,747	464	4,352,619
Hispanic	329	4,316,353	461	460,738

* Note: Counts include all youth 12 to 18 years old. These are the ages included in the 2003 report of findings. However, only those youth 12½ to 18 years of age are included in the analyses described in the present report. Also, the numbers in this table may differ slightly from those given in the main report due to corrections made after the weighting process was completed.

Table A-20. Unweighted and weighted counts of responding youth in third followup waves for selected subgroups

	Third followup (Round 4)			
	Wave 8		Wave 9	
Subgroup	Unwtd.	Wtd.	Unwtd.	Wtd.
Total 12½-18 years*	1,966	27,036,977	3,020	26,957,342
12½ to 13	398	6,455,570	769	6,345,363
14 to 18	1,568	20,581,407	2,251	20,611,979
Male	1,048	13,648,559	1,545	13,810,108
Female	918	13,388,419	1,475	13,147,234
Non-Black, non-Hispanic	1,403	18,377,337	2,142	18,403,593
Black, non-Hispanic	258	4,210,915	439	4,220,186
Hispanic	305	4,448,725	439	4,333,563

* Note: Counts include all youth 12½ to 18 years old (i.e., the ages included in the analyses described in the present report). The numbers in this table may differ slightly from those given in the main report due to corrections made after the weighting process was completed.

Table A-21. Estimated coverage rates for selected subgroups of youth by wave

Subgroup	Wave 1 Coverage rate	Wave 2 Coverage rate	Wave 3 Coverage rate
Male	0.71	0.68	0.65
Female	0.68	0.69	0.65
Race/Ethnicity:			
Non-Hispanic white, other non-Hispanic	0.69	0.69	0.65
Non-Hispanic Black	0.69	0.67	0.63
Hispanic	0.74	0.66	0.62
Age Group			
9 to 11	0.70	0.69	0.70
12 to 13	0.74	0.71	0.75
14 to 18	0.67	0.67	0.57

A.2.6 Calculation of Dyad Weights

A dyad is defined to be a unique youth-parent combination. Parents had to complete question F4 or a later question in order for the questionnaire to be considered complete. Parents who were too ill to complete the questionnaire, physically or mentally impaired, or could only communicate in a language other than English or Spanish were considered ineligible in Waves 1 through 3. Parents who were no longer living with the sampled youth or who were physically or mentally disabled were considered to be ineligible for the followup waves. The response rates achieved in the study for the sampled parents are documented in Appendix B. Note that a sampled parent may be associated with up to two dyads in NSPY.

A responding dyad is one in which both the youth and associated parent were eligible for NSPY and each completed their respective interviews. A nonresponding dyad is one in which both the youth and parent were eligible for NSPY, but one or the other did not complete the interview. All other dyads were considered to be ineligible. For example, dyads consisting of youth who were not eligible for NSPY (such as youth in group quarters, institutionalized youth, non-English/Spanish speaking youth)

were considered to be ineligible dyads. Youth for whom a parent interview was not required (such as emancipated youth living alone or married youth) are by definition "nondyads" and were also considered to be ineligible.

Thus, the dyad weight is essentially a youth weight that represents a subset of the youth population, namely, youth living with their parents (or other specified caregiver). The initial weights assigned to dyads were the youth weights defined in Section A.2.3. These initial dyad weights were adjusted for nonresponse using the same auxiliary variables used for youth nonresponse adjustment. Again, the MART software was used to define appropriate weighting cells within which the nonresponse adjustments were applied. Finally, the nonresponse-adjusted dyad weights were poststratified to CPS control totals using the raking algorithm described in Section A.2.5. The same raking dimensions used to calibrate the youth weights were used to calibrate the dyad weights. The final weighted counts of responding dyads are summarized in Tables A-22 through A-25 by wave.

Table A-22. Unweighted and weighted counts of responding dyads in recruitment waves for selected subgroups

	Recruitment waves (Round 1)					
	Wave 1		Wave 2		Wave 3	
Subgroup	Unwtd.	Wtd.	Unwtd.	Wtd.	Unwtd.	Wtd.
Total 9-18 years*	3,106	39,716,564	2,209	39,757,285	2,305	40,171,280
9 to 11	1,033	12,544,155	869	12,698,100	924	12,833,899
12 to 13	989	7,913,378	616	8,020,827	683	8,056,295
14 to 18	1,084	19,259,031	724	19,038,358	698	19,281,086
Male	1,632	20,321,184	1,129	20,473,904	1,191	20,650,392
Female	1,474	19,395,381	1,080	19,283,380	1,114	19,520,888
Non-Black, non-Hispanic	2,193	27,810,424	1,571	27,485,831	1,616	27,860,879
Black, non-Hispanic	454	6,116,262	346	6,303,599	342	6,228,104
Hispanic	459	5,789,879	292	5,967,855	347	6,082,297

* Note: Counts include all youth 9 to 18 years old and their parents (i.e., dyads). These are the ages included in the first through third semiannual reports. However, only those youth 12½ to 18 years of age and their parents are included in the analyses described in the present report. Also, the numbers in this table may differ slightly from those given in the main report due to corrections made after the weighting process was completed.

Table A-23. Unweighted and weighted counts of responding dyads in first followup waves for selected subgroups

	First followup (Round 2)			
	Wave 4		Wave 5	
Subgroup	Unwtd.	Wtd.	Unwtd.	Wtd.
Total 12-18 years*	1,945	27,711,435	2,934	27,614,850
12 to 13	635	8,197,998	1,165	8,237,424
14 to 18	1,310	19,513,437	1,769	19,377,426
Male	1,036	14,186,400	1,470	14,145,558
Female	909	13,525,035	1,464	13,469,291
Non-Black, non-Hispanic	1,396	19,284,631	2,117	19,209,401
Black, non-Hispanic	262	4,364,886	428	4,364,275
Hispanic	287	4,061,917	389	4,041,174

* Note: Counts include all youth 12 to 18 years old and their parents (i.e., dyads). These are the ages included in the fourth and fifth semiannual reports. However, only those youth 12½ to 18 years of age and their parents are included in the analyses described in the present report. Also, the numbers in this table may differ slightly from those given in the main report due to corrections made after the weighting process was completed.

Table A-24. Unweighted and weighted counts of responding dyads in second followup waves for selected subgroups

	Second followup (Round 3)			
	Wave 6		Wave 7	
Subgroup	Unwtd.	Wtd.	Unwtd.	Wtd.
Total 12-18 years*	2,032	27,911,524	3,055	28,261,862
12 to 13	592	8,428,634	1,089	8,446,595
14 to 18	1,440	19,482,890	1,966	19,815,267
Male	1,089	14,370,242	1,551	14,469,914
Female	943	13,541,283	1,504	13,791,947
Non-Black, non-Hispanic	1,471	19,186,657	2,165	19,355,140
Black, non-Hispanic	264	4,490,451	455	4,335,865
Hispanic	297	4,234,417	435	4,570,857

* Note: Counts include all youth 12 to 18 years old and their parents (i.e., dyads). These are the ages included in the 2003 report of findings. However, only those youth 12½ to 18 years of age and their parents are included in the analyses described in the present report. Also, the numbers in this table may differ slightly from those given in the main report due to corrections made after the weighting process was completed.

Table A-25. Unweighted and weighted counts of responding dyads in third followup waves for selected subgroups

	Third followup (Round 4)			
	Wave 8		Wave 9	
Subgroup	Unwtd.	Wtd.	Unwtd.	Wtd.
Total 12½-18 years*	1,869	26,747,263	2,862	26,494,922
12½ to 13	384	6,441,631	743	6,308,711
14 to 18	1,485	20,305,632	2,119	20,186,211
Male	1,007	13,567,985	1,467	13,625,718
Female	862	13,179,279	1,395	12,869,205
Non-Black, non-Hispanic	1,337	18,117,476	2,021	18,055,895
Black, non-Hispanic	250	4,210,915	431	4,187,706
Hispanic	282	4,418,872	410	4,251,321

* Note: Counts include all youth 12½ to 18 years old (i.e., the ages included in the analyses described in the present report) and their parents (i.e., dyads). The numbers in this table may differ slightly from those given in the main report due to corrections made after the weighting process was completed.

In previous reports based on earlier waves of the NSPY, a separate set of parent weights was constructed in addition to the dyad weights described above (e.g., see Hornik et al., 2000). These parent weights were used to compute descriptive statistics of parent characteristics, such as those given in Chapter 3 of this report. The parent associational analyses, such as those given in Chapter 6, were computed for youth-parent dyads, using the dyad weights defined at the beginning of this section. For the fourth and final round of NSPY, the calculation of separate parent weights was complicated by the fact that the population of inference was restricted to youth 12½ to 18 years of age. In order to derive the required parent weights, it would have been necessary to develop household population control totals corresponding to this restricted age group for youth (e.g., see Hornik et al., 2000 for a description of the parent weighting methodology). Data from the Current Population Survey were used for this purpose previously for the youth age range of 12 to 18 years of age, but could not be used for the 12½- to 18-year age range because age in the CPS public use files is reported only in single years of age. Because of the main focus on dyad rather than parent-level analyses and the fact that construction of parent weights would have been problematic, a decision was made to replace the parent-level analyses that had been included in previous reports with the corresponding dyad-level analyses.

An examination of the effect of this change using previous waves of NSPY data indicated that it had very little impact on specific and general measures of TV and radio ad recall. For example, using the parent weights, 38.9 percent of parents in Waves 1 through 7 reported seeing parent-targeted Campaign TV ads at least once a week, as compared with 38.7 percent of youth-parent dyads with a parent who reported seeing parent-targeted Campaign TV ads at least once a week. Similarly, 50.0 percent of parents in Waves 1 through 7 reported having seen or heard general TV or radio ads at least weekly, compared to 50.4 percent of youth-parent dyads. Some additional comparisons for exposure-related measures are given below.

- Across Waves 1 through 7, the mean number of Campaign ad viewings per month was 8.46 for parents and 8.40 for dyads.
- Across Waves 1 through 7, 33.3 percent of parents reported no viewings per month of the "parenting skills/personal efficacy" platform ads, compared to 34.4 percent of dyads.
- Across Waves 1 through 7, 21.3 percent of parents reported seeing newspapers/magazine ads at least weekly, compared with 22.2 percent of dyads.

The effect of the change was slightly greater for other types of characteristics collected in the parent interview, but did not change the general conclusions. For example, the change in parent internet use between 2000 and 2003 using the parent weights was 13.3 percent compared to a 15.7 percent change using the dyad weights. Both of these change estimates were statistically significant at the 95 percent confidence level. Estimates of change of parents' recall of stories about drug use also tended to be slightly larger for the dyad-level analysis than for the parent-level analysis. For example, the change between 2000 and 2003 in the percent of parents that reported having noticed stories on TV or radio news programs dealing with drug use was –5.1 percent, compared with –5.8 percent of dyads. Some additional comparisons for these other types of variables are given below.

- Between 2000 and 2003, the change in the percent of parents who reported having noticed stories on TV talk shows/news magazine programs dealing with drug use was 0.9 percent, compared with –0.8 percent of dyads. However, despite the change in sign of the year-to-year change, neither change estimate was statistically significant.
- Between 2000 and 2003, the change in the percent of parents who reported having noticed stories on non-news radio programs dealing with drug use was 4.6 percent, compared with 3.8 percent of dyads.
- Between 2000 and 2003, the change in the percent of parents who reported hearing at least one weekly story with a drug theme in at least one medium in the past 12 months was –3.2 percent, compared with –3.8 percent of dyads.

In summary, the results obtained using the parent and dyad weights are generally very similar. Although there are instances where one set of weights yielded a significant result but the other did not, the main findings were not affected by the choice of weights.

A.2.7 Calculation of Longitudinal Weights

Starting with Wave 4, longitudinal weights were constructed for the purpose of analyzing the delayed effects of current exposures on future outcomes. Separately for youth and dyads, seven sets of longitudinal weights were constructed, one for each of the longitudinal samples listed below:

(1) Youth (dyads) completing both Wave 1 and Wave 4 interviews;

(2) Youth (dyads) completing both Wave 2 and Wave 5 interviews;

(3) Youth (dyads) completing both Wave 3 and Wave 5 interviews;

(4) Youth (dyads) completing both Wave 4 and Wave 6 interviews;

(5) Youth (dyads) completing both Wave 5 and Wave 7 interviews;

(6) Youth (dyads) completing both Wave 6 and Wave 8 interviews; and

(7) Youth (dyads) completing both Wave 7 and Wave 9 interviews.

To construct the required longitudinal weights, eligible youth and dyads in a given longitudinal sample were assigned initial weights equal to the final cross-sectional (raked) weights developed for the corresponding "baseline" wave. For example, the initial weights for the first of the samples listed above were the final Wave 1 raked weights. Similarly the initial weights for the second of the samples listed above were the final Wave 2 raked weights, and so on. These initial weights were then adjusted for nonresponse in the subsequent followup wave using the same general methods used to construct the cross-sectional weights. Along with the auxiliary variables described in Section A.2.4, selected variables derived from the baseline interview were used to form the nonresponse adjustment cells. These included use of marijuana in the prior wave, exposure to the Campaign in the prior wave as measured by the general exposure score, prior anti-marijuana beliefs/attitude index, and prior marijuana social norm index.

In addition to youth weights, longitudinal weights for analysis of dyads were also developed. The procedures used to create the dyad weights were analogous to those used to create the youth weights. Tables A-26 and A-27 summarize results of the youth nonresponse adjustment process for each of the seven longitudinal samples described above. Tables A-28 and A-29 give the corresponding results for dyads. Note that the longitudinal weights are used to derive the counterfactual projection (CFP) weights for delayed effects analysis (see Appendix C for details).

A.2.8 Calculation of Replicate Weights for Variance Estimation

The sampling errors of estimates derived from NSPY data have been calculated using a replication technique. This replication method uses 100 replicates to measure the variance of the full-sample estimates. The method developed for NSPY reflects the variance due to each stage of sampling (i.e., selecting PSUs, segments within PSUs, dwelling units within segments, and persons within dwelling units), as well as finite population correction factors at both the PSU and segment levels. Full technical details of the method developed for NSPY are given in Rizzo and Judkins (2004).

Table A-26. Summary of youth nonresponse adjustment process for longitudinal samples with a recruitment wave as the initial wave

Youth[1] NR adjustment	Sample with recruitment wave as initial wave		
	Wave 1 to 4	Wave 2 to 5	Wave 3 to 5
Weighted count of sampled youth[2]	24,457,867	29,093,085	29,003,981
Weighted count of responding youth[2]	22,737,155	25,089,561	26,028,645
Weighted (conditional) response rate[2, 3]	93.0%	86.2%	89.7%
Top 5 MART predictors[4]	URBAN_C	URBAN_C	PARRAC1
	MJATTBEL	PARRAC1	URBAN_C
	MJNORM	YUTHRAC1	MULTIDU
	PCT3_37	MJATTBEL	YUTHRAC1
	W4IMPAGE	MJNORM	EXTRFAM
Number of NR (MART) cells formed	10	4	6
Range of adjustment	1.00-1.40	1.04-1.40	1.03-1.30

[1] Youth 12 to 18 years of age at followup.

[2] Weights are the final youth (raked) weights derived in the prior wave as defined by formula (8). Only those originally-sampled youth in households for which eligibility was ascertained in Round 2 (Waves 4 and 5) are included in the weighted counts. The percentage of households for which eligibility was ascertained in Round 2 was generally lower for Wave 4 than for Wave 5.

[3] Conditional on responding youth refielded for followup.

[4] URBAN_C indicates the type of area: 0 = rural, 1 = town or exurban area, 2 = second city of metropolitan area, 3 = suburb of metropolitan area, 4 = urban part of metropolitan area.

MULTIDU indicates the multifamily status of dwelling unit: 0 = single family, 1 = first unit of a multifamily dwelling, 2 = all other units of a multifamily dwelling, blank = not available. For Wave 1, a blank indicates that DU is part of the building permit sample.

W4IMPAGE indicates age of youth at followup.

YUTHRAC1 indicates race/ethnicity of youth.

PARRAC1 indicates the race/ethnicity of parent.

EXTRFAM indicates whether other family members live in household.

MJATTBEL is the anti-marijuana attitudes/belief index at prior wave.

MJNORM is the marijuana social norm index at prior wave.

PCT3_37 = tercile rank of the segment with respect to percent of persons in labor force who are employed.

Table A-27. Summary of youth nonresponse adjustment process for longitudinal samples with a followup wave as the initial wave

Youth[1] NR adjustment	Sample with followup wave as initial wave			
	Wave 4 to 6	Wave 5 to 7	Wave 6 to 8	Wave 7 to 9
Wtd. count of sampled youth[2]	28,323,047	28,491,618	27,959,748	27,585,109
Wtd. count of responding youth[2]	26,348,086	26,027,366	25,864,233	25,782,185
Wtd. (conditional) resp. rate[2, 3]	93.0%	91.4%	92.5%	93.5%
Top 5 MART predictors[4]	URBAN_C	URBAN_C	URBAN_C	URBAN_C
	NONREL	RACEETHP	YGEIORD	MJNORM
	MULTIDU	MULTIDU	MJNORM	MJATTBEL
	MJATTBEL	MJATTBEL	MJATTBEL	RACEETHC
	RACEETHP	MJNORM	INT_AGEC	RACEETHP
Number of NR (MART) cells	6	9	3	9
Range of adjustment	1.03-1.35	1.01-1.35	1.03-1.38	1.02-1.24

[1] Youth 12 to 18 years of age at followup in Waves 6 and 7; youth 12½ to 18 years of age at followup in Waves 8 and 9.

[2] Weights are the final youth (raked) weights derived in the prior wave as defined by formula (8).

[3] Conditional on responding youth refielded for followup.

[4] URBAN_C indicates the type of area: 0 = rural, 1 = town or exurban area, 2 = second city of metropolitan area, 3 = suburb of metropolitan area, 4 = urban part of metropolitan area.

MULTIDU indicates the multifamily status of dwelling unit: 0 = single family, 1 = first unit of a multifamily dwelling, 2 = all other units of a multifamily dwelling, blank = not available. For Wave 1, a blank indicates that DU is part of the building permit sample.

RACEETHP indicates the race/ethnicity of parent.

INTAGEC indicates age of youth at followup.

RACEETHC indicates race/ethnicity of youth.

NONREL indicates whether or not an unrelated person also lives in the household.

MJATTBEL is the anti-marijuana attitudes/belief index at prior wave.

MJNORM is the marijuana social norm index at prior wave.

YGEIORD is the general exposure score at prior wave.

Table A-28. Summary of dyad nonresponse adjustment process for longitudinal samples with a recruitment wave as the initial wave

Youth[1] NR adjustment	Sample with recruitment wave as initial wave		
	Wave 1 to 4	Wave 2 to 5	Wave 3 to 5
Weighted count of sampled youth[2]	23,733,188	27,579,278	27,906,065
Weighted count of responding youth[2]	21,574,713	23,784,110	24,984,815
Weighted (conditional) response rate[2, 3]	90.9%	86.2%	89.5%
Top 5 MART predictors[4]	URBAN_C	IMPAGEP	IMPAGEP
	W4IMPAGP	MJATTBEL	URBAN_C
	PARRAC1	URBAN_C	YUTHRAC1
	MJATTBEL	YUTHRAC1	MJATTBEL
	W4IMPAGC	NONPARS	PARRAC1
Number of NR (MART) cells formed	7	6	2
Range of adjustment	1.00-1.36	1.00-1.26	1.03-1.13

[1] Youth 12 to 18 years of age at followup.

[2] Weights are the final dyad (raked) weights derived in the prior wave using formulas analogous to formula (8). Only those originally-sampled youth in households for which eligibility was ascertained in Round 2 (Waves 4 and 5) are included in the weighted counts. The percentage of households for which eligibility was ascertained in Round 2 was generally lower for Wave 4 than for Wave 5.

[3] Conditional on responding youth refielded for followup.

[4] URBAN_C indicates the type of area: 0 = rural, 1 = town or exurban area, 2 = second city of metropolitan area, 3 = suburb of metropolitan area, 4 = urban part of metropolitan area.

W4IMPAGC indicates age of youth at followup.

YUTHRAC1 indicates race/ethnicity of youth.

W4IMPAGP and IMPAGEP indicate the age of parent.

PARRAC1 indicates the race/ethnicity of parent.

NONPARS indicates that no parents live in household.

MJATTBEL is the anti-marijuana attitudes/belief index at prior wave.

Table A-29. Summary of dyad nonresponse adjustment process for longitudinal samples with a followup wave as the initial wave

Youth[1] NR adjustment	Sample with followup wave as initial wave			
	Wave 4 to 6	Wave 5 to 7	Wave 6 to 8	Wave 7 to 9
Wtd. count of sampled youth[2]	27,538,070	27,388,403	27,161,610	26,301,002
Wtd. count of responding youth[2]	25,358,483	24,614,881	24,958,538	24,137,751
Wtd. (conditional) resp. rate[2, 3]	92.1%	89.9%	91.9%	91.8%
Top 5 MART predictors[4]	URBAN_C	INT_AGEP	MJNORM	INT_AGEP
	INT_AGEP	URBAN_C	URBAN_C	URBAN_C
	RACEETHP	RACEETHC	MJATTBEL	RACEETHP
	MJATTBEL	RACEETHP	MULTIDU	MJNORM
	PCT3_49	GENDERP	INT_AGEP	EXTRFAM
Number of NR (MART) cells	5	6	8	6
Range of adjustment	1.04-1.46	1.02-1.35	1.03-1.42	1.04-1.47

[1] Youth 12 to 18 years of age at followup in Waves 6 and 7; youth 12½ to 18 years of age at followup in Waves 8 and 9.

[2] Weights are the final dyad (raked) weights derived in the prior wave using formulas analogous to formula (8).

[3] Conditional on responding youth refielded for followup.

[4] URBAN_C indicates the type of area: 0 = rural, 1 = town or exurban area, 2 = second city of metropolitan area, 3 = suburb of metropolitan area, 4 = urban part of metropolitan area.

MULTIDU indicates the multifamily status of dwelling unit: 0 = single family, 1 = first unit of a multifamily dwelling, 2 = all other units of a multifamily dwelling, blank = not available. For Wave 1, a blank indicates that DU is part of the building permit sample.

RACEETHC indicates race/ethnicity of youth.

INT_AGEP indicates the age of parent.

RACEETHP indicates the race/ethnicity of parent.

GENDERP indicates sex of parent.

EXTRFAM indicates whether other family members live in household.

MJATTBEL is the anti-marijuana attitudes/belief index at prior wave.

MJNORM is the marijuana social norm index at prior wave.

PCT3_49 = tercile rank of the segment with respect to percent of vacant housing units for occasional use.

To permit the calculation of sampling errors, a series of replicate weights were calculated and included in the data files. The first step was to construct 100 replicates as described in Rizzo and Judkins (2004). After the replicates were created, the full set of weight adjustment procedures described in the previous sections was applied to each replicate. This meant that each set of replicate weights was adjusted for nonresponse and raked to specified CPS-based control totals. Thus, the calculated replicate weights reflect all of the adjustments used to create the full-sample weights. As a result of this process, each youth and dyad record in the analysis file has 101 different weights, an overall weight and 100 replicate weights.

A.3 Confidence Intervals, Data Suppression, and Design Effects

Ninety-five percent confidence intervals have been provided for every statistic in the Detail Tables. These intervals indicate the margin of error due to sampling. If the same general sampling procedures were repeated a large number of times, and a statistic of interest and its confidence interval were recalculated for each of those independent samples, the true population value would be contained

within 95 percent of the calculated confidence intervals. The confidence intervals reflect the effects of sampling and of the adjustments that were made to the weights. They do not generally reflect measurement error in the questionnaires. Details on how the confidence intervals were calculated are given in Section A.3.1.

Some estimates were suppressed if they were subject to large sampling errors. This was determined on the basis of both the sample size and the width of the confidence interval associated with the estimate. For example, estimated proportions near 0 or 1 based on small samples were more likely to be suppressed than other estimates. The criteria for suppression are summarized in Section A.3.2.

Finally, estimates of design effects for selected statistics are summarized in Section A.3.3 for reference. The design effect is defined to be the ratio of the variance of an estimate derived from the NSPY sample design to the corresponding variance that would have been obtained from a hypothetical simple random sample of the same size.

A.3.1 Confidence Intervals

Variances were estimated for NSPY using a replication approach. Using the replicate weights described in Section A.2.8, estimates were computed for each set of replicate weights, and these replicate-specific estimates were then used to generate the required variances. For example, let $\bar{y}$ denote the estimated mean of a survey item of interest based on the full sample, and let $\bar{y}_k$ denote the corresponding estimate for replicate k. Then the standard error of $\bar{y}$ was computed as:

$$SE(\bar{y}) = \sqrt{\sum_{k=1}^{100} h_k(\bar{y}_k - \bar{y})^2} \qquad (9)$$

where h_k is a specified constant for the kth replicate (e.g., see Rizzo and Judkins, 2004).

Once the standard error estimates were obtained, confidence intervals were calculated using approximations similar to those developed for the National Household Survey on Substance Abuse (NHSDA). For means of continuous variables, the confidence intervals were formed by assuming that the sample statistic had a t-distribution with 100 degrees of freedom (i.e., the number of replicates). In the NHSDA, it was assumed that the sample statistic had a normal distribution. That was equivalent to assuming a t-distribution with an infinite number of degrees of freedom. The 100 degrees of freedom used for NSPY estimates was expected to be slightly more conservative. The standard error was multiplied by 1.98 instead of 1.96 to form a 95 percent confidence interval as follows:

$$\text{lower bound} = \bar{y} - 1.98\ SE(\bar{y}) \text{ and upper bound} = \bar{y} + 1.98\ SE(\bar{y}).$$

For confidence intervals around proportions, it is assumed that a logistic transform of the estimated proportion, $\hat{p}$, has a normal distribution. This results in confidence limits that are strictly between 0 and 1. The formulas for the corresponding lower and upper confidence limits are:

$$\text{lower bound} = \frac{1}{1 + \exp\left\{-\left[\log\left(\frac{\hat{p}}{1-\hat{p}}\right) - 1.98\frac{\sqrt{\text{var}(\hat{p})}}{\hat{p}(1-\hat{p})}\right]\right\}}$$

and

$$\text{upper bound} = \frac{1}{1 + \exp\left\{-\left[\log\left(\frac{\hat{p}}{1-\hat{p}}\right) + 1.98\frac{\sqrt{\text{var}(\hat{p})}}{\hat{p}(1-\hat{p})}\right]\right\}}.$$

For example, if an estimated percentage is 0.5 percent with a standard error of 0.4 percent, rather than calculating a confidence interval of -0.3 percent to +1.3 percent using the usual normal (or t-based) approximations, the logistic formulas above yield a confidence interval of 0.1 percent to 2.4 percent.

Estimated proportions equal to 0 or 1 pose additional difficulties for variance estimation and calculation of confidence intervals. The calculated standard error of 0 is not meaningful for such estimates. The following approximation was used instead to obtain a confidence interval associated with an estimated proportion of 0:

$$\text{lower bound} = 0 \text{ and upper bound} = \frac{2F_{2,n}^{-1}(1-\alpha/2)}{n + 2F_{2,n}^{-1}(1-\alpha/2)}$$

where $F_{2,n}^{-1}(1-\alpha/2)$ is the 1-$\alpha/2$ quantile of an F distribution with 2 and n degrees of freedom (Korn and Graubard, 1999), and n refers to the effective sample size defined to be the actual sample size divided by an average design effect (as suggested by D. Judkins and P. Zador). An average design effect of 2 was used to calculate the confidence intervals (e.g., see Section A.3.3. for examples of design effects).

For an estimated proportion of 1, the corresponding confidence interval was calculated as

$$\text{lower bound} = \frac{nF_{n,2}^{-1}(\alpha/2)}{2 + nF_{n,2}^{-1}(\alpha/2)} \text{ and upper bound} = 1.$$

For example, suppose that the sample size for a given domain is 500. Then, using the above formulas, the upper confidence limit on an estimate of 0 percent would be 1.5 percent, while the lower confidence limit on an estimate of 100 percent would be 98.5 percent.

For differences between proportions where one of the estimates is 0 or 1, a slight modification of the above formulas was used. The approximation used for a confidence interval around an estimated difference of 0 was:

$$\text{lower bound} = \frac{-2F_{2,n}^{-1}(1-\alpha/2)}{n + 2F_{2,n}^{-1}(1-\alpha/2)} \text{ and upper bound} = \frac{2F_{2,n}^{-1}(1-\alpha/2)}{n + 2F_{2,n}^{-1}(1-\alpha/2)},$$

where $F_{2,n}^{-1}(1-\alpha/2)$ is the 1-$\alpha/2$ quantile of an F distribution with 2 and n degrees of freedom, and n was estimated as the harmonic average of the two sample sizes. For a difference of proportions where only one of the estimates was zero, the standard error of the nonzero estimate was used to impute the standard error of the zero estimate, and the corresponding confidence interval was then computed in the usual way.

A.3.2 Suppression

Estimates that were based on very small sample sizes or were otherwise subject to very large sampling errors were suppressed. For example, if an estimate was based on a sample size of only two youth, and the estimated proportion of them who had a particular characteristic was 50 percent, then the corresponding confidence interval would be 5.7 percent to 94.3 percent, which is too wide to be meaningful.

Several suppression rules were used in this report The primary rule was to suppress any estimate based on an effective sample size of 30 or less, where the effective sample size for a statistic was calculated as the nominal sample size divided by the design effect.

Estimated proportions between 0 and 0.5 were also suppressed if

$$\frac{\sqrt{\text{var}(\hat{p})}}{\hat{p}\log(1/\hat{p})} > 0.225.$$

Similarly, estimated proportions between 0.5 and 1.0 were suppressed if

$$\frac{\sqrt{\text{var}(\hat{p})}}{(1-\hat{p})\log(1/(1-\hat{p}))} > 0.225\ .$$

Note that the above rules meant that larger effective sample sizes were required to avoid suppression of estimated proportions approaching 0 or 1.

Estimated proportions equal to 0 or 1 were suppressed if the effective sample size for the domain was 140 or less. This corresponds to confidence limits of 0.000 to 0.026 on an estimate of 0 and 0.974 to 1.000 on an estimate of 1.

A.3.3 Design Effects and Effective Sample Sizes

The design effect is defined to be the ratio of the variance of an estimate derived from a complex sample design to the corresponding (hypothetical) variance that would have been achieved if a simple random sample of the same size had been used. The effective sample size is defined to be the nominal sample size divided by the design effect. For complex multistage sample designs such as that employed for the NSPY, design effects exceeding 1.0 can be expected. Such design effects result primarily from the use of clustering (by PSUs and segments) and varying sampling rates in the NSPY sample design. Tables A-30 (youth) and A-31 (dyads) summarize estimated design effects and the corresponding effective sample sizes for selected statistics. The results in the tables are intended to illustrate the range of design effects to be expected for statistics derived from the NSPY by age group and wave. In Table A-30, the statistics for which design effects were computed include: (a) the mean attitudes-and-belief-about marijuana-use index, (b) the mean perceptions-of-social-norms-about-marijuana-use index; and (c) the proportion of youth who have ever used marijuana. Table A-31 is similar except that the statistics refer to youth-parent dyads. As can be seen in the tables, the design effects vary considerably from statistic to statistic, partially reflecting true differences but also reflecting the instability (variance) of the estimated design effects. In general, design effects for the total sample tend to be larger than for age subgroups because of the increased variation in sampling weights associated with the total sample.

Table A-30. Design effects and effective sample sizes for selected youth statistics by age group

Wave/ Statistic	12½ to 13*		14 to 15		16 to 18		Total 12½ to 18	
	Effective samp.size	Design effect	Effective samp.size	Design effect	Effective samp.size	Design effect	Effective samp.size	Design effect
WAVE 1								
Anti-MJ index	331	2.38	430	1.28	469	1.30	1,200	1.62
MJ norm index	603	1.31	400	1.38	544	1.12	1,712	1.14
Ever used MJ	500	1.58	432	1.27	379	1.59	992	1.96
WAVE 2								
Anti-MJ index	396	1.24	222	1.78	286	1.35	868	1.47
MJ norm index	414	1.19	237	1.66	247	1.57	705	1.80
Ever used MJ	330	1.49	167	2.36	232	1.66	805	1.58
WAVE 3								
Anti-MJ index	343	1.58	289	1.29	232	1.64	576	2.25
MJ norm index	313	1.74	258	1.45	234	1.62	561	2.31
Ever used MJ	492	1.10	224	1.66	236	1.60	529	2.44
WAVE 4								
Anti-MJ index	425	1.16	597	1.34	454	1.26	1,372	1.36
MJ norm index	560	0.88	537	1.50	455	1.26	1,522	1.23
Ever used MJ	354	1.39	539	1.49	355	1.62	1,019	1.83
WAVE 5								
Anti-MJ index	675	1.35	636	1.57	577	1.46	1,462	1.88
MJ norm index	768	1.19	557	1.79	556	1.52	1,564	1.76
Ever used MJ	776	1.17	444	2.25	681	1.24	1,786	1.54
WAVE 6								
Anti-MJ index	349	1.31	579	1.37	551	1.33	1,163	1.70
MJ norm index	439	1.04	514	1.54	467	1.56	1,359	1.46
Ever used MJ	403	1.13	546	1.45	366	1.99	910	2.17
WAVE 7								
Anti-MJ index	876	0.92	754	1.41	694	1.29	1,887	1.47
MJ norm index	798	1.01	763	1.40	643	1.39	1,684	1.64
Ever used MJ	661	1.29	1,249	0.91	729	1.28	1,940	1.51
WAVE 8								
Anti-MJ index	280	1.41	834	0.77	714	1.30	1,413	1.39
MJ norm index	329	1.20	613	1.04	594	1.57	1,603	1.23
Ever used MJ	330	1.20	610	1.05	712	1.31	1,541	1.28
WAVE 9								
Anti-MJ index	1,064	0.72	785	1.36	934	1.26	2,213	1.36
MJ norm index	589	1.31	746	1.44	913	1.29	2,169	1.39
Ever used MJ	368	2.09	879	1.22	783	1.51	1,944	1.55

* Due to the aging of the original recruitment samples, only youth 12½ to 18 years of age are included in the present analysis.

Table A-31. Design effects and effective sample sizes for selected dyad statistics

Wave/youth age group	12½ to 13*		14 to 15		16 to 18		Total 12½ to 18	
	Effective samp.size	Design effect	Effective samp.size	Design effect	Effective samp.size	Design effect	Effective samp.size	Design effect
WAVE 1								
Anti-MJ index	329	2.27	362	1.44	461	1.22	1,032	1.77
MJ norm index	586	1.27	403	1.29	509	1.11	1,523	1.20
Ever used MJ	461	1.62	398	1.31	344	1.63	944	1.94
WAVE 2								
Anti-MJ index	361	1.26	237	1.56	238	1.49	729	1.62
MJ norm index	348	1.30	272	1.36	227	1.56	661	1.78
Ever used MJ	308	1.47	167	2.21	215	1.65	697	1.69
WAVE 3								
Anti-MJ index	324	1.60	260	1.36	226	1.52	577	2.10
MJ norm index	284	1.82	223	1.58	218	1.57	547	2.22
Ever used MJ	448	1.16	205	1.71	239	1.43	590	2.06
WAVE 4								
Anti-MJ index	438	1.09	536	1.41	393	1.36	1,216	1.46
MJ norm index	599	0.80	468	1.62	378	1.42	1,281	1.38
Ever used MJ	338	1.42	496	1.52	299	1.79	841	2.11
WAVE 5								
Anti-MJ index	657	1.33	576	1.67	551	1.43	1,483	1.77
MJ norm index	775	1.13	519	1.85	505	1.56	1,455	1.80
Ever used MJ	810	1.08	425	2.25	600	1.32	1,724	1.52
WAVE 6								
Anti-MJ index	315	1.38	569	1.33	480	1.42	1,149	1.63
MJ norm index	383	1.14	517	1.47	406	1.68	1,324	1.42
Ever used MJ	325	1.34	513	1.48	327	2.08	826	2.27
WAVE 7								
Anti-MJ index	787	0.99	764	1.34	700	1.19	1,857	1.42
MJ norm index	755	1.03	779	1.31	659	1.26	1,720	1.53
Ever used MJ	697	1.18	1,152	0.95	726	1.20	1,843	1.51
WAVE 8								
Anti-MJ index	264	1.45	726	0.84	671	1.31	1,330	1.41
MJ norm index	318	1.20	610	1.00	617	1.42	1,608	1.16
Ever used MJ	283	1.35	560	1.09	631	1.39	1,439	1.30
WAVE 9								
Anti-MJ index	1,054	0.70	782	1.31	829	1.32	2,097	1.36
MJ norm index	585	1.27	739	1.39	926	1.18	2,025	1.41
Ever used MJ	397	1.87	811	1.26	763	1.44	1,908	1.50

* Due to the aging of the original recruitment samples, only youth 12½ to 18 years of age are included in the present analysis.

References

Durbin, J. (1967). Design of multi-stage surveys for the estimation of sampling errors. *Applied Statistics*, 16, 152-164.

Friedman, J. (1999). *Tutorial: Getting Started with MART in Splus*. Stanford University.

Judkins, D., Cadell, D.M., and Sczerba, K. (2000). *Costs and benefits of a permit sample late in the decade*. Presented at the Annual Meeting of the American Statistical Association.

Korn, E.L. and Graubard, B.I. (1999). *Analysis of health surveys*. New York: Wiley.

Rizzo, L. and Judkins, D. (2004) *Replicate variance estimation for the national survey of parents and youth. Expected to be published in Proceedings of the Survey Methods Research Section, American Statistical Association.*

Miller, D.R. and Hodges, K. (1994). *A population density approach to incorporating an urban-rural dimension into small area lifestyle clusters*. Presented at the Annual Meeting of the Population Association of America.

Weiss, M.J. (2000). *The clustered world: How we live, what we buy, and what it all means about who we are*. Boston: Little, Brown and Co.