# Appendices

# Appendix A

# Sample Design, Development of Weights, Confidence Intervals and Data Suppression, and Geography

This appendix provides a more detailed discussion of the same points discussed in Chapter 2 of this report. However, it is still a condensed discussion. A more detailed report on the sampling plan is available as Chapter 2 of the overall Evaluation Plan from the National Institute on Drug Abuse (NIDA). This appendix is separated into four main sections along the lines suggested by the title.

## A.1 Sample Design

The youth and their parents were found by door-to-door screening of a scientifically selected sample of about 34,700 dwelling units for Wave 1, 23,000 dwelling units for Wave 2, and 23,300 for Wave 3. These dwelling units were spread across about 1,300 neighborhoods in 90 primary sampling units (PSUs) for Wave 1 and about 800 neighborhoods each, in the same primary sampling units for Waves 2 and 3. 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 housing were included in the sample. Youth living in institutions, group homes, and dormitories were excluded.

For subsequent followup waves there has been no new selection of dwelling units or of youth. However, an original sampled parent could be replaced by a newly selected parent if the original selected parent were no longer eligible.

The sampling was arranged to get adequate numbers of youth in each of three targeted age ranges: 9 to 11, 12 to 13, and 14 to 18. These age ranges were judged to be important analytically for evaluating the impact of the Media Campaign. Within households with multiple eligible youth, up to two youth were selected during the three initial recruitment waves.

Parents were defined to include natural parents, adoptive parents, and foster parents who lived in the same household as the sample youth. Stepparents were also usually treated the same as parents unless they had lived with the child for less than 6 months. When there were no parents present, an adult caregiver was usually identified and interviewed in the same manner as actual parents. No absentee parents were selected. During the three initial recruitment waves, when more than one parent or caregiver was present, one was randomly selected. No preference was given to selecting mothers over fathers. Parents or caregivers of both genders were selected at equal rates. This was done to be able to measure the impact of the Media Campaign separately on mothers and fathers. During the subsequent followup waves, the most knowledgeable parent was selected if the original sample parent was no longer eligible (e.g., no longer living with child at least two nights a week, or mentally or physically

disabled). When there were two sample youth who were not siblings living in the same household, a parent figure was selected for each.

The following discussion about sample selection is divided into two major subsections. The first describes the selection of the screening sample and the second describes the selection of youth and parents. As indicated earlier, all of the major sampling activities occurred during Waves 1 through 3, i.e., the three initial recruitment waves). The sample for Wave 4 was a subset of youth and parents selected for Wave 1 that included all Wave 1 respondents plus a small subsample of Wave 1 nonrespondents (see Section A.1.3 for details).

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

The screening sample was selected using a dual-frame, multistage design. One frame was of housing built by late 1991 as listed by Westat in a sample of areas using field personnel and maps. This frame was called the *area frame*. The second frame consisted of building permits issued for new housing between January 1990 and December 1998. The dual-frame approach was used to improve survey reliability. By sampling new construction from permits, it was possible to spread the sample out more evenly, which resulted in improved reliability (Judkins, Cadell, and Sczerba, 2000). Housing units built in 1990 and 1991 had two chances of selection since they appeared in both frames. To correct for this overcoverage, the screening questionnaire in Waves 1 through 3 instructed the interviewers to ask the age of the housing for a sample selected from the area frame. Any housing units in the area frame built after April 1, 1990, were ineligible for the survey. Housing units built in the first 3 months of 1990 were kept under the assumption that there was some lag between the issuance of a permit and the construction of the building. Housing units built after 1998 had no chance of selection in either frame. Also, a housing unit had no chance of selection if built during the 1990s in jurisdictions where no permit was required. Finally, modular housing built during the 1990s was inadvertently omitted from the permit sample. These three factors implied a household coverage rate of about 98 percent.

New mobile homes placed on sites between 1991 and 2000 had a chance of selection through the missed mobile home procedure. This worked as follows. In a sample of segments (as defined below), interviewers were instructed to canvas the segment on their first visit for mobile homes 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.

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

The area screening sample was selected in three stages. The first stage consisted of selecting a sample of PSUs. The PSUs were generally metropolitan areas and groups of nonmetropolitan counties. The second stage consisted of segments. Each segment was a block or group of contiguous blocks with a minimum housing unit count in 1990 of about 60. The third stage consisted of individual dwelling units.

## PSU Selection

The PSUs were selected from a design stratified by region, metropolitan status, per capita income, percentage minority population, and PSU size. The National Survey of Parents and Youth (NSPY) PSUs were drawn as a subset of Westat's 1991 master sample. This master sample comprised 100 PSUs. Of these, 90 were selected for NSPY. One reason for using a subset of these 100 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 a prior survey, thereby reducing the area sampling costs.

The following paragraphs describe how the 100-PSU master sample was drawn and how it was subsampled for NSPY use. The PSUs in the underlying frame were constructed using 1990 Decennial Census information based on the following general criteria:

- Each PSU consisted of a single county, a group of counties, or a metropolitan statistical area (MSA).
- The PSUs were geographically contiguous, mutually exclusive, and covered the United States.
- Nonmetropolitan PSUs did not cross state boundaries.
- Each PSU had at least 15,000 total population as of 1990.
- Each PSU was designed to be as easily traversable by an interviewer or lister as possible given population density, minimum size constraints, and natural topography.

This constructed frame included 1,404 PSUs, with no PSU having a 1990 population larger than 5,400,000 (the New York, Chicago, and Los Angeles PMSAs were divided into three, two, and two PSUs, respectively). From this constructed frame, 100 PSUs were selected in 1991 for the master sample.

The 100-PSU master sample was selected using probability-proportionate-to-size (PPS) sampling with 1990 population as a measure of size. Twenty-four PSUs with populations greater than 2,100,000 were certainty selections (selected with probability 1). The remaining 1,380 PSUs were assigned to 38 strata for PSU selection. These strata were defined to satisfy the following criteria:

- Each stratum represented a 1990 population of roughly 4 to 5 million persons.
- The 38 strata were nested within eight primary strata defined by census region (Northeast, South, Midwest, and West) and PSU metropolitan/nonmetropolitan status.
- The strata within each primary stratum were constructed to be heterogeneous in PSU population size (for metropolitan primary strata), per capita income, and percentage minority population.

Using the Durbin-Brewer method (Durbin, 1967), 76 PSUs were sampled from the 38 strata (two PSUs per stratum) with probability proportionate to their 1990 population.

The NSPY PSU sample was a random subsample of 90 PSUs from the 100-PSU master sample. The noncertainty strata were grouped into superstrata. One stratum was then selected from each superstratum. Within the selected stratum, one of the two sample PSUs was randomly deselected. In order to eliminate 10 PSUs, 10 superstrata were formed, each with the same number of strata. The

superstrata were formed from the 38 noncertainty strata and two pairs of small certainty PSUs. This yielded an even four strata per superstratum. Each superstratum contained eight sample PSUs, each of which represented a population of approximately 2.1 million people. One PSU was dropped from each superstratum for a total of 10 eliminated PSUs, as required.

In forming the superstrata, there was some grouping of strata across regions because not every region had a number of strata that was a multiple of four and higher priority was given to avoiding grouping across metropolitan status. This approach was expected to increase the variance of regional estimates. To counteract this increased variance, a special set of weights was built for regional analyses. For this special set of weights (developed solely for cross-sectional analyses of Waves 1 through 3 data), the probabilities of retention associated with the superstrata were ignored and, instead, the PSUs in each region were weighted by metropolitan status up to the total population reported in those areas in 1990. This approach reduced variance on regional analyses but increased bias and variances for other statistics. Therefore, the regional weights were used only for regional analyses in Waves 1 through 3.

## Area Segment Selection

NSPY segments consisted of groups of neighboring blocks with a minimum count of 60 dwelling units in the 1990 Census. By using blocks instead of larger units of geography, such as tracts or official block groups, the size of the listing task was reduced. However, some blocks had very small and even zero populations. These were collapsed to meet the minimum requirement of 60 dwelling units. A total of 1,180 such segments were selected for Wave 1. The sample segment counts were smaller for Waves 2 and 3 with 689 segments selected for Wave 2 and 694 segments for Wave 3. For the Wave 2 and 3 segments, all dwelling units were screened for date of construction. On average, approximately 27 dwelling units per segment were sampled in Wave 1 with a slightly larger average of 29 dwelling units per segment in Waves 2 and 3. The large minimum size of 60 dwelling units was designed to avoid selecting adjacent neighbors for the sample. This had the advantage of reducing contamination of interviews by prior interviews in neighboring houses.

The segments for Wave 1 were a subset of segments originally selected and listed for another survey in late 1991. (The listing process consisted of sending field workers out to every segment. Using a map of the segment, the field worker prepared a list of dwelling units within the segment.) In addition to saving the cost of a new listing of 1,180 segments, the use of these old listings had the advantage of eliminating most housing built during the 1990s. This might have been a drawback for another survey, but the NSPY had a separate sample of building permits to cover 1990s construction. Any dwelling units built in the 1990s in area segments had to be screened out, so using an old list actually made the total data collection more efficient. The segments for Waves 2 and 3 were from the same 1991 frame but were listed in a separate process 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 1999 for the stratum that the PSU represented. From the earlier survey, there was a total of 2,065 segments available. These segments had been selected in a systematic PPS fashion,[1] where the measure of size counted African American and Hispanic households more heavily than other households. This approach resulted in an oversample of segments with strong concentrations of minority population. This oversample was not desired for NSPY. Since just 1,180 of the 2,065 segments were required, the segments were subsampled with probabilities such that overall probability

[1] A systematic PPS selection is one where the frame is systematically sorted and then an unequal probability sample is drawn with PPS. The systematic sorting induces a set of joint probabilities of selection that minimizes the total variance.

of selection became proportional to total households without any special emphasis on minority households. This was done by using a measure of size (MOS) that was proportional to the ratio of desired overall probability to the original probability:

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

## Dwelling Unit Selection in Area Segments

As mentioned above, the 1,180 segments for Wave 1 had been listed by contractor staff in late 1991 and early 1992. These lists of housing addresses were keyed. From the keyed files, a systematic PPS sample was drawn with a fixed national target of 30,993 dwelling units. (When combined with the permit sample of 3,407 newly built dwelling units, the total initial sample size was 34,400.) The measure of size was defined as the weight for the segment so that the final dwelling unit sample would be closer to an equi-probability sample (i.e., a sample in which every dwelling unit had the same chance of selection). These 30,993 dwelling units were split into two release groups by segment, with about 590 segments in each release group. For Wave 2, the 689 segments were supplemented with 2,875 new construction dwelling units for a total of 23,000 dwelling units. All of the Wave 2 segments were listed in the fall of 1999. For Wave 3, the 694 segments were supplemented with a permit sample of 3,052 for a total of 23,300 units.

For a subsample of the sample dwelling units, there was a quality control check on the original 1991/1992 listing. For all single-family housing, the interviewer checked for hidden apartments (such as converted basements, garages, and attics) that might have been missed by the lister. Any detected hidden apartments were added to the sample. Also, in a subsample of multifamily housing structures, the interviewer checked for missed apartments. Using these procedures, 192 missed dwelling units were added to the sample. Also, as mentioned above, there was a check for new mobile homes. This procedure added 99 sample mobile homes to the sample. Thus the combined sample from area segments was 31,284 dwelling units. Because the Waves 2 and 3 segments were listed in the fall of 1999, this process was not employed for these waves.

## Selection of the Permit Screening Sample

A separate building permit sample was drawn for the three initial waves of NSPY to prevent problems caused by outdated information on block sizes. The data collection procedures for selecting the area segment involved sampling with PPS in the 1990 Census. PPS sampling with 1990 data strongly reduced between-segment variation to the extent that there was a strong correlation between total population in 1990 and eligible population in 1999. New construction would weaken that correlation. To avoid the potentially high between-segment variance caused by a weakened correlation, only pre-1990 census housing from the area segments were interviewed. This was accomplished 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. A separate sample of postcensus housing was drawn from a frame of building permits. This procedure was introduced at the U.S. Census Bureau in the 1960s and continues to be used for all major household surveys conducted by it. It is used at Westat for large surveys conducted late in a decade.

Permit sampling was possible because most localities required that a permit be obtained before building a residential structure and because the U.S. Census Bureau conducted a regular census of permit activity. This census of local governments has been conducted every month for active offices

and annually for less active offices. A benefit of the census has been that it could be used to select specific offices and months from which to draw an efficient sample of permits for national estimates.

The stages of permit sampling were similar to those in the area frame, but there were five instead of three. First, only permits issued within the 90 sample PSUs were selected. Next, a sample of building permit offices (BPOs) was selected. These were the local county and city offices that issue building permits and keep records about them. At the third stage, a sample of segments was selected, where a segment was defined to be the set of permits issued by an office within a specific time frame. At the fourth stage, individual permits were selected. After selection of the permits, a lister visited all the building sites for the selected permits to list all the housing units that were found there. After listing of housing units within sample segments, the final sample of dwelling units was selected.

The total dwelling unit sample size from the permit frame was set so that the proportion of the total sample selected through the permit frame would roughly equal the proportion of the total 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 dwelling unit sample size from the permit frame for Wave 1 was 3,407, equal to about 10 percent of the total initial sample. In Wave 2 the dwelling unit sample from the permit frame was 2,875 units compared to 20,125 area sample dwelling units for Wave 2. Because the permit frame covered only until the end of 1998, there was no coverage of new housing units that were permitted and built in 1999 or in 2000. The Wave 3 permit sample was 3,052 units while the area sample for Wave 3 consisted of 20,248 units. For Wave 3, there was no coverage of new housing units that were permitted and built in 1999, 2000, and the first half of 2001.

### A.1.2 Selection of Youth and Parents (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 households selected as a result of subsampling, one youth was selected from each age range represented, but no more than a total of two youth were selected. The parents and caregivers for the sample youth were than identified and one was randomly selected. The practice of sampling up to two youth when any are selected 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 means that youth in the less rare age domains were sampled at a higher rate if they happen to have a sibling in a rarer age domain. Similar procedures have been used successfully on other surveys. This approach was particularly advantageous for NSPY because the precision requirements for parents were specified in terms of the youth age domains. A mother with children in two or three of the age ranges would be counted toward the parent precision targets for each range in which one of her children was selected. Thus, concentrating the youth selections in a smaller set of households generated a more efficient parent sample. This approach also increased the amount of directly collected sibling data. On the negative side, it increased design effects slightly for older youth, but this had been anticipated and was counteracted by using a slightly larger nominal sample size for this age range.

To carry out this sampling efficiently, it was convenient to divide eligible households into three strata based on the combination of ages represented by the youth in the household. Because youth aged 12 to 13 were the rarest age domain, households containing such youth were always selected. They are 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. For Waves 2 and 3, there were no subsampling within either stratum. Finally, 14- to 18-year-olds represented the most common age domain and were most sharply subsampled so that they constituted a third stratum. Thus, the following strata were used:

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

Table A-A shows estimates of the youth population by stratum from Wave 1 of NSPY. These estimates were prepared using final Wave 1 NSPY youth weights. They were broadly consistent with earlier estimates obtained from the Current Population Survey (CPS). The retention rates represent the percentage of the screened households of the given type that were retained in NSPY. These rates applied only to Wave 1. The retention rates for Waves 2 and 3 were modified slightly.

**Table A-A. 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 | 3,431,546 | 7,998,814 | 4,094,726 | 15,525,086 |
| At least one 9- to 11-yr.-old but no 12- to 13-yr.-olds | 70% | 8,449,930 | 9,102,823 | 0 | 3,100,064 | 12,202,887 |
| At least one 14- to 18-yr.-old but no 9- to 13 -yr.-olds | 45% | 9,545,207 | 0 | 0 | 11,862,093 | 11,862,093 |
| Total | | 25,766,069 | 12,534,369 | 7,998,814 | 19,056,883 | 39,590,066 |

The mechanics of sample selection then worked as follows. When DUs were selected from the area and permit segments, they were randomly assigned to one of three sampling rules:

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

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

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

For sampling rule *A*, the interviewer was instructed to induct the household into the sample only if it contained a youth aged 12 or 13. For sampling rule *AB,* the interviewer inducted the household into the sample if it contained one or more youth aged 9 to 13. For sampling rule *ABC,* the interviewer inducted 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.

Eligibility rates have been estimated based on the results from the initial recruitment waves. Table A-B shows eligibility rates in Wave 1 for households assigned to the different screener groups. These rates are lower than were predicted based on CPS tabulations (also shown in Table A-B). There was

significant undercoverage at Wave 1—on the order of 30 percent undercoverage. The reasons for this undercoverage are not known but persisted in Waves 2 and 3.

**Table A-B. Wave 1 eligibility rates**

| Screener group | Screener sample (%) | Wave 1 age eligibility rate (%) | CPS predictions of eligibility rates (%) |
|---|---|---|---|
| A | 30.1% | 05.6% | 07.5% |
| AB | 24.9% | 10.8% | 15.2% |
| ABC | 45.0% | 19.9% | 24.4% |
| Total | 100.0% | 12.2% | 17.0% |

For Waves 2 and 3, stratum B was sampled at the same rate as 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. Because there would be no new sample after Wave 3, the only way to achieve an oversample of 12- to 13-year-olds after Wave 3 was to oversample the 9- to 11-year-olds at Waves 2 and 3.

For the followup waves, the sample will become older because the 9-year-olds are not replenished. Several plans for replenishing the sample of 9-year-olds were considered but they ran into serious operational problems. The most serious problem was that about 37 percent of 8-year-olds have older siblings. To give a chance of selection to these 8-year-olds when they turn 9, a third youth would have to be sampled in many households. That would have resulted in a serious change in existing data structures. There were also lesser problems with sampling and tracking 8-year-olds who did not have older siblings. Given the low level of attention that the Media Campaign was paying to 9- to 11-year-olds, it did not seem worth the high cost to maintain a large sample of children aged 9 to 11 past Wave 3.

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

The only chance of selection for mobile homes was through the area frame because the permit frame did not cover these DUs. Therefore, the screener instructed the interviewer to skip the year-built question for mobile homes. This procedure was efficient for all but Wave 1. The 1991 listings used for these waves included all trailer sites occupied in 1991 but missed all new trailer parks and all isolated mobile homes parked in new locations. To provide coverage of these mobile homes, interviewers recanvassed a subsample of the segments for mobile homes. Any segment from which the first listed DU was selected was marked for the special canvass. Any mobile homes were compared with the old listing sheets to see whether they were enumerated. All previously unenumerated mobile homes were added to the sample in these segments for Wave 1. This procedure yielded a sample of 99 missed mobile homes for Wave 1.

Another activity that took place during the screening process for Wave 1 was called the *missed 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 of the segment to determine whether the apartment was listed. If missed by the lister, 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 lister. Any previously unlisted apartments were added to the screening sample. At Wave 1, this procedure generated a sample of 192 missed DUs.

The missed mobile home and missed DU procedures were not used for Waves 2 and 3. The listings used for those waves were prepared in mid-1999, making them fairly fresh for interviewing in late 2000 and early 2001. Because of the screening and sampling procedures, all stick and modular housing built after 1998 was excluded from the sampling frame. In addition, all mobile homes placed after the listing period in mid-1999 had no chance of selection.

### A.1.2.1 Youth and Parent Selection (Waves 1 through 3)

The procedure for Waves 1 through 3 was to prepare a list of eligible youth in each sample household and sample one youth within each nonempty age range, subject to a maximum of two sample youth per household. In a household with youth in all three of the age ranges, one youth from the 12-to-13 range was selected. A random decision was then made to either select a second youth from the 9-to-11 range or from the 14-to-18 range. Within an age range, all youth had the same probability of selection. At least one and no more than two youth were selected for every sample household. The interviewers then determined the relationship of all adults in the household to each sample youth and the relationship between the two sample youth if two were selected. If two sample youth were siblings (whole, half, or step), the computer selected one adult from the set of adults in the household who were classified as a parent or caregiver of either youth. If two nonsiblings (such as cousins) 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 questionnaire.

During Waves 1 through 3, a random parent instead of the most knowledgeable or cooperative parent was selected for several reasons. Most importantly, parent statistics were to be prepared in addition to youth statistics. Because the most knowledgeable and cooperative parent in two-parent households is often the mother, a nonrandom selection would have resulted in a sample consisting mostly of mothers with very little data on fathers. To be able to measure the penetration of the Media Campaign with fathers as well as mothers, random selection of parents was used for Waves 1 through 3.

Parents were defined as biological, adoptive, step, or foster parents sharing a roof with a youth. Caregivers were defined as persons serving *in loco parentis* for youth who did not live with their parents. Some distinctions were made between these categories for sampling purposes. Stepparents were considered parents for sampling purposes only if they had lived with their stepchild for at least 6 months. In addition, the exact nature of the relationship between the adult and the youth were recorded for analytic purposes. Henceforth, in this discussion, the term *parent* will be used to refer to both parents and caregivers unless otherwise specified.

In multifamily households, all youth within an age range were given an equal chance of selection. If two selected youth were cousins or are not related at all (as in the case of a live-in nanny with her own children), a separate parent was selected for each family with a sample youth.

For youth with divorced or separated parents, priority was given to the household where the youth spent the majority of the year. Only these households were eligible for selection. The only parent figure eligible for selection was the natural/adoptive parent with whom the youth spent most of the year and any stepparent present in that household. It was possible to select the stepparent without selecting the natural/adoptive parent.

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 who lived in the same DU as the youth 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 above, youth residing in group quarters were not sampled during the recruitment phase; youth living in boarding schools and college dormitories were, therefore, 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 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, youth who live at home or in private apartments while attending college were sampled. 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. This special exclusion of dormitory residents did pose some special challenges to the weighting process. To poststratify the sample, it was necessary to estimate the dormitory population from the 1990 decennial census and then to carry that estimate forward, in order to subtract it from more current CPS estimates of the entire noninstitutional population aged 9 to 18.

One complication of the dormitory exclusion concerned the length of the field period. For example, Wave 2 started in July 2000. To maintain a stable sampling universe throughout the interviewing period, youth who were currently living in boarding schools and dormitories or who were expected to be in those living arrangements by the end of the wave were excluded. Note that this had the effect of excluding from the spring wave high school seniors who were planning to live in dormitories in the fall. Note that this applied only in the initial recruitment wave. In the subsequent followup waves, such youth were excluded only if they lived in a dormitory or boarding school at the time of initial screening (not any time during data collection).

---

[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.

Table A-C shows the counts of interviewed youth at Wave 1 by age and by household stratum. Within households completing the household roster, person-level response rates were high. For example, in Wave 1, extended interviews were obtained for 88 percent of sampled parents and 91 percent of sampled youth. Appendix B provides additional details on response rates.

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

| Household composition | Rostered households | Parents | Youth per age domain | | | |
|---|---|---|---|---|---|---|
| | | | 9-11 | 12-13 | 14-18 | Total 9-18 |
| At least one 12- to 13-yr.-old | 1,191 | 1,057 | 340 | 1,080 | 389 | 1,809 |
| At least one 9- to 11-yr.-old but no 12- to 13-yr.-olds | 826 | 733 | 749 | 0 | 230 | 979 |
| At least one 14- to 18-yr.-old but no 9- to 13 -yr.-olds | 584 | 503 | 0 | 0 | 524 | 524 |
| Total | 2,601 | 2,293 | 1,089 | 1,080 | 1,143 | 3,312 |

### A.1.3 Selection of Followup Sample for Wave 4

Under the NSPY sample design, subsamples of youth and parents selected for the initial recruitment waves (i.e., Waves 1 through 3) will be retained for followup in subsequent data collection waves. No new samples will be 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, that is, those parents and youth who were selected but who did not complete a Wave 1 interview were refielded in Wave 4. The "extra" youth and parents that were obtained in Wave 4 were used only for cross-sectional analyses at Wave 4. Appendix B provides details on response rates.

## 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 3, there were six sets of final weights in all, three for national analyses and three for regional analyses. There were national weights for youth, for parents, and for youth-parent dyads. These repeated for regional analyses. For Wave 4, separate regional weights were not prepared. Instead, in addition to national cross sectional weights, two sets of longitudinal weights were created, one for lagged analysis and one for stable analysis. These weights were used to reflect selection probabilities and to compensate for nonresponse and undercoverage. The adjustments for undercoverage involved a process called raking. In the raking process, the weights were adjusted in such a manner that the sums of weights for important domains agreed with those from independent more reliable sources. The final weight for a respondent, including nonresponse adjustments and raking, can be viewed as the number of population members that each respondent represented. Details about the weighting process are given in the following sections.

## A.2.1 Baseweights

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

Calculation of the baseweights was done by considering the probability of selection at each stage: PSU, segment, dwelling unit, and person. The calculation of these probabilities at each stage was fairly straightforward. However, since the person selection could be carried out only in households where the screener was completed, the person-level baseweight also reflected nonresponse adjustment and, in the case of the parent weights, an adjustment for household undercoverage.

For Waves 1 through 3, the baseweight for a dwelling unit is generally

$$BW_{DUi} = \frac{1}{\Pr\{\text{PSU}\}\Pr\{\text{segment} \mid \text{PSU}\}\Pr\{\text{DU} \mid \text{segment}\}}.$$

For permit segments, there were also some adjustments for failure to find the permits for a particular segment and for the lack of coverage of new housing in jurisdictions where building permits were not required. These adjustments were based on statistics from the Census Bureau's reports on construction starts. Also, in Wave 2, the BPO weights were trimmed to avoid inflating the variances.

These dwelling unit-level baseweights were then adjusted for screener nonresponse as discussed in Section A.2.3 below. After adjustment for screener nonresponse, the adjusted weight was further adjusted for screener-based subsampling. Dwelling units in Wave 1 had been preassigned to three screening groups: *A*, *AB*, and *ABC*. However, for Waves 2 and 3 dwelling units were assigned only to screening groups *AB* and *ABC*. Dwelling units in the *A* screening group were retained in sample only if there was a youth aged 12 to 13 present in the dwelling unit. Dwelling units in the *AB* screening group were retained in sample only if there was a youth aged 9 to 13 present. Dwelling units in the *ABC* screening group were retained in sample only if there was a youth aged 9 to 18 present. These rules were developed as a means to efficiently oversample dwelling units containing youth aged 12 to 13 and (to a lesser extent) those containing youth aged 9 to 11. Based on these screening rules, all dwelling units in all waves with youth aged 12 to 13 were retained with certainty so no adjustment was required to their weights. Also in Waves 2 and 3, those dwelling units with a youth aged 9 to 11 present, but no youth aged 12 to 13, were retained with certainty so again no adjustment was required to their weights. However, in Waves 2 and 3, those dwelling units with a youth aged 9 to 11 present, but no youth aged 12 to 13, had a probability of retention of 0.7, so their weights were adjusted upward by a factor of 1.4286. Similarly, those dwelling units with a youth aged 14 to 18 present, but none aged 9 to 13, had a probability of retention of just 0.45, so their weights were adjusted upward by a factor of 2.2222.

After this stage in the calculation, different paths were taken for the calculation of youth and parent baseweights. However, from this point on, the procedures for Waves 1 through 3 were the same. The youth path is described first.

There were three age classes for youth sampling purposes: 9 to 11, 12 to 13, and 14 to 18. If there were youth present in all three age ranges, the first step in youth subsampling was to select two out the three

age ranges. The 12-to-13 range was always selected with certainty. One of the other two was selected with equal probability. So the first component in the youth probability of selection for youth aged 9 to 11, or 14 to 18 in such households was a factor of 0.5. Next, within each sample age range, one youth was selected from however many were present. For example, if there were 4 youth present in an age range, the probability of selection within the range was 0.25. The two factors were multiplied together to create a youth within-household probability of selection. The youth baseweight was then calculated as the quotient of the adjusted baseweight for the household divided by the within-household probability of selection for the youth.

The parental probability of selection was more complex. In simple nuclear families, the probability of selection for a parent was simply 1.0 for single-parent households and 0.5 for two-parent households, but a variety of other living arrangements were encountered. Some households contained nephews and nieces of the householder where the householder or his/her spouse was reported as the caregiver for the nephew or niece, but not both were so reported. Sometimes, one or two parents of the nephew or niece were present. Sometimes a grandparent was considered the caregiver of the nephew or niece. Other households contain couples who was not married but each had their own children. Some households contain boarders, housekeepers, or nannies who had their own children present.

When one youth was selected, a random parent/caregiver was selected from the set of parents and caregivers for that youth. When two siblings were selected, a random parent/caregiver was selected from the set of parents and caregivers identified for either sibling. When two youth were selected who were not siblings, one parent/caregiver was selected from the "pool" of parents and caregivers for each. If these pools overlapped, it might still be the case that just one parent figure was selected; thus, the parent's probabilities of selection depended on their relationship to the youth in the household. While the relationship of every adult in the household was established to the sample children, this information was not collected about nonsample children. These relationship data were imputed using the available data about household composition. Each parent and caregiver's probability of selection was then computed over all possible youth samples from the household.

Given the complexity of the parent/caregiver concept for NSPY, it was realized that no post-stratification or raking to independent estimates of parents would be possible. In order to correct for undercoverage despite the lack of ability to perform such adjustment, the decision was made to rake the household weights prior to applying the within-household probabilities of selection for parents. This raking is discussed below in Section A.2.4.

For Wave 4, the starting point for the weighting process was the set of sampling weights derived for Wave 1. Because no new youth were selected in Wave 4, the weights from Wave 1 were used as the base weights for youth in Wave 4. These weights were nonresponse adjusted and then raked to the youth population totals at Wave 4. For originally selected parents, Wave 1 weights were used as the base weights for Wave 4. It was possible to select a new parent if the originally selected parent was no longer eligible, for example, in the case of a divorce. In this case the newly selected parent was treated as a substitute for the originally selected parent.

A new feature in Wave 4 was the construction of longitudinal weights. Youth and dyads who were eligible in Wave 1 and were still eligible in Wave 4 were given longitudinal weights that were based on the Wave 1 weights. There was no new raking on the longitudinal weights since these weights were intended to estimate the longitudinal attributes of the Wave 1 population. However, these weights were nonresponse adjusted using the same methods as the cross-sectional weights.

## A.2.2 Nonresponse Adjustments

In general, it was hoped that there were groups of households where the decision to respond to a survey was unrelated to substantive characteristics of interest such as substance abuse. Complex modeling techniques were employed to find groups of households with different response rates. The variables that were available to define such groups were mostly from the 1990 Decennial Census and described the block groups containing the households. Within a group, the weighted response rate was calculated. The baseweight was then divided by the group response rate to obtain the nonresponse-adjusted weight for a household. Households in groups with low response rates received large upward adjustments in their weights. Intuitively, this meant that those hard-to-reach households that were interviewed despite being hard to reach ended up receiving larger weights than households that were easy to reach. If the groups were formed well, this procedure could eliminate nonresponse bias. If too many were formed, however, the variation in weights caused by groups with low response rates could hurt survey reliability.

The goal was to develop procedures that would form enough but not too many groups. To this end, special software was created (built on top of data mining software) to form the groups. A set of about 60 household characteristics was used in conjunction with the special software. Some examples of the characteristics used include local percentages of persons in certain age groups, persons of certain race and ethnicity, homeowners versus renters, persons in mobile homes, U.S. citizens versus noncitizens, and persons with incomes below the poverty level.

This type of adjustment was done separately for the doorstep and roster phases of the screener, for youth nonresponse, for parent nonresponse, and for dyad nonresponse.

### A.2.2.1 Screener Nonresponse Adjustment

This adjustment was done in two phases and applied only to Waves 1 through 3. The first phase was to adjust for doorstep nonresponse where it was never determined whether eligible youth were present at the address. The second phase was to adjust for roster nonresponse where it was known that the household did contain eligible youth, but it was not possible to prepare a household roster and select a sample of youth and parents.

In the doorstep phase, a dwelling unit was considered to be a respondent if information about the presence of children had been collected from either the occupants of the household or from their neighbors. In addition, if the dwelling unit was selected in an area segment and was not a mobile home, information on the age of the structure was required in order to be considered a complete doorstep screener. As mentioned in Appendix B, the screener response rate was 95.1 percent for Wave 1, 95.7 percent for Wave 2, and 95.5 percent for Wave 3. The adjustment factors for screener nonresponse varied from 1.0 to 1.7 for both Waves 1 and 2 and the factors varied from 1.0 to 1.6 for Wave 3.

In the roster phase, an eligible household was considered to be a respondent if an adult resident of the household had been found who was willing to provide a roster of the occupants of the household, their ages, and their relationships to the sample children. If any of this information was withheld, it was impossible to select the youth and parent sample so the household was classified as a nonrespondent. As mentioned in Appendix B, the roster response rate was 74.4 percent for Wave 1,

74.6 percent for Wave 2, and 75.3 percent for Wave 3. The adjustment factors for roster nonresponse varied from 1.1 to 1.6 for both Waves 1 and 2, but the factors varied from 1.1 to 1.7 for Wave 3.

### A.2.2.2 Youth

Youth who answered D13 or any subsequent question were considered respondents. This was the last question on general ad exposure prior to prompting their recall with a display of several real advertisements. Nonrespondents included those whose parents refused consent or otherwise failed to provide consent, those who refused personal assent, and those who were just never reached to do the interview for any reason. Among those who did not complete the questionnaire, a difference was drawn between those who physically or mentally were incapable of completing it 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 only in that the weight of ineligible youth was not redistributed to responding youth through the nonresponse adjustment. Included in the category of ineligible youth were those who could not communicate in English or Spanish. Since the television and radio components of the Media Campaign were only in these languages, it seemed appropriate to classify those who cannot communicate in either language as ineligible for the evaluation. Also potentially included in the ineligible youth category were young people who stepped into parental roles for other youth aged 9 to 18. This might occur by reason of marrying an older person with such youth or by reason of caring for younger siblings.

The set of the same 60 household characteristics used for doorstep and roster nonresponse adjustment, as well as some additional characteristics, were used in conjunction with special adjustment software to develop an appropriate set of response cells for all sampled eligible youth. The additional characteristics included items such as whether both of the youth's parents were in the household, whether the youth was an only child, the total number of youth living in the household, and whether there was a nonrelative living in the household. All of these variables were obtained from the household roster. The resulting set of response cells was then used to adjust the weights of the respondents at the youth level. As mentioned in Appendix B, the youth response rate was 90.6 percent for Wave 1, 91.6 percent for Wave 2, 90.9 percent for Wave 3, and 93.6 percent for Wave 4. The adjustment factors for youth nonresponse varied from 1.0 to 1.5 for Wave 1, from 1.1 to 1.7 to for Wave 2, from 1.0 to 1.6 for Wave 3, and from 1.0 to 1.4 for Wave 4.

Note that for Wave 4, both cross-sectional and longitudinal weights were derived for analysis. The two sets of weights differ slightly because for cross-sectional analysis, a respondent was defined to be a sampled youth who completed the Wave 4 interview, whether or not the Wave 1 interview was completed; whereas for longitudinal analysis, a respondent was defined to be a youth who completed both Wave 1 and Wave 4 interviews. In Wave 4, about 94 percent of the eligible youth who completed the Wave 1 interview were longitudinal responders. For longitudinal youth nonresponse adjustment, the adjustment factors ranged from 1.0 to 1.6.

### A.2.2.3 Parent

The parent nonresponse adjustment procedure was very similar to that for youth. 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. As mentioned in Appendix B, the parent response rate was 88.4 percent for Wave 1, 87.6 percent for Wave 2, 86.9 percent for Wave 3, and 90.3 percent for Wave 4. The adjustment factors for parent nonresponse varied from 1.0 to 1.5 for Wave 1, from 1.0 to 1.7 for Wave 2, from 1.1 to 1.7 for Wave 3, and from 1.0 to 1.5 for Wave 4.

#### A.2.2.4 Youth-Parent Dyads

Respondents for this analysis were defined as youth who responded and whose parents also responded to the survey. Therefore, both the youth and the parent had to be eligible and have completed their respective surveys to count as a respondent. Nonrespondents included all eligible nonresponding youth, but also included any youth who may have responded but whose parent did not. Youth who were not eligible for the youth weights were also not eligible for dyad analysis. Youth who did not have a corresponding sampled parent interviewed (such as emancipated youth or married youth) were considered ineligible for this set of weights. Also, youth who were eligible and completed an interview but whose parents were ineligible were considered ineligible for the Youth-Parent dyad weights.

The same characteristics used for youth nonresponse adjustment were used for dyad nonresponse adjustment. Again, the special adjustment software was implemented to define appropriate nonresponse adjustment cells, and weighting adjustments were computed using that set of cells. The dyad response rate was 85.7 percent for Wave 1, 86.4 percent for Wave 2, 85.7 percent for Wave 3, and 89.6 for Wave 4. The adjustment factors for dyad nonresponse varied from 1.1 to 1.6 for Wave 1, from 1.1 to 1.5 for Wave 2, from 1.1 to 1.6 for Wave 3, and from 1.0 to 1.5 for Wave 4.

In addition to cross-sectional weights, longitudinal dyad weights were also developed for Wave 4. Among eligible responding dyads in Wave 1, 91.4 percent were longitudinal responders (i.e., also responded in Wave 4). For longitudinal nonresponse adjustment the factors ranged from 1.0 to 1.04 for Wave 4.

### A.2.3 Raking

Raking is a commonly used procedure in which survey estimates are controlled to marginal population totals. In theory, the estimates should differ from the population values only as a result of sampling error. In practice, other error sources such as residual nonresponse and coverage errors still may have an important effect on the accuracy of the estimates. The goal of raking is to reduce biases due to undercoverage and nonresponse, and to reduce the sampling error of the estimates. Raking may be thought of as an iterative form of poststratification, in which the weights are consecutively ratio-adjusted to multiple sets of control totals until the resulting weights converge to the control totals in each dimension. The sample sizes of the marginal distributions are the important determinants of the stability of the raking procedure, not the cells formed by a complete cross-classification of the variables. This permits the use of more auxiliary variables or control totals than in poststratification. For this reason we chose to rake the household, youth, and dyad weights rather than poststratify them. However, when sample sizes permitted, some raking dimensions were defined by crossing two variables to preserve the correlation structure in the data.

The parent weights were not raked because no control totals exist for parents as defined by the NSPY. However, estimates of total households with youth between the ages of 9 and 18 were available from the January 2000 CPS for Wave 1, and for Wave 2 the October 2000 CPS data were available.

Wave 3 used the average of March 2001 and April 2001 CPS data. For Wave 3, this average centered the control totals in the middle of the data collection period. For Wave 4 a regression line was fit to a year of CPS data and the estimate for October of 2001 was used as the population total. Marginal household control totals were obtained from the CPS for the following four raking dimensions:

- Household Race/Ethnicity (Non-Hispanic-white + Other Non-Hispanic, Non-Hispanic-Black, Hispanic) by Presence of Male Age 28 or Older in the Household (Yes/No);
- Youth Age Group Composition of Household (any age 12 to 13 present, age 9 to 11 present but no age 12 to 13, age 14 to 18 present but no age 9 to 13);
- Household Race/Ethnicity (Non-Hispanic-white, Non-Hispanic-Black, Other Non-Hispanic, Hispanic); and
- Census Region (Northeast, Midwest, South, West).

After the household doorstep and roster nonresponse adjustments, the household weights were raked to the first three sets of control totals to produce the household weights that were used in creating national parent baseweights. The household weights were raked again on all four dimensions for use in creating regional parent baseweights. Convergence was obtained in Wave 1 after three iterations for the national household weights and six iterations for the regional weights. Convergence was obtained in Waves 2 and 3 after four iterations for the national household weights and six iterations for the regional weights.

For youth, estimates of the total age 9 to 18 civilian population were also obtained from the January 2000 CPS and October 2000 CPS for Wave 1 and Wave 2, respectively. As with the household totals, the youth totals for Wave 3 were based on the average of March 2001 and April 2001 CPS data. From these control totals the civilian noninstitutional group quarters population was excluded, as estimated from the 1990 Census Public Use Micro-data System (PUMS) files. Marginal control totals were obtained for the categories defined by the three raking dimensions:

- Gender (M, F) x Age Group (ages 9 to 11, 12 to 13, and 14 to 18);
- Race/Ethnicity (Non-Hispanic-white, Non-Hispanic-Black, Other Non-Hispanic, Hispanic) x Age Group (ages 9 to 11, 12 to 13, and 14 to 18); and
- Census Region (Northeast, Midwest, South, West) x Age Group (ages 9 to 11, 12 to 13, and 14 to 18).

After the Youth and Youth-dyad nonresponse adjustments, both sets of weights were raked to the first two sets of control totals to produce the final national youth and Youth-dyad weights for use in analysis. Both sets of nonresponse-adjusted weights were raked again on all three dimensions to create regional weights for use in making regional estimates. Convergence was obtained after four iterations for the national weights in Waves 1 and 2, but converged in three iterations for the national weights in Wave 3. Convergence was obtained in six iterations for the regional weights for Waves 1 through 3.

Coverage rates are given in Table A-D for youth by age, race, and gender. The coverage rate was calculated as the ratio of the sum of the weights before raking to the control total. Coverage rates were not computed for Wave 4, because the Wave 4 sample was a subset of the Wave 1 sample.

**Table A-D**
**Coverage rates**

| 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 | | | |
| 9-11 | 0.70 | 0.69 | 0.70 |
| 12-13 | 0.74 | 0.71 | 0.75 |
| 14-18 | 0.67 | 0.67 | 0.57 |

## A.3 Confidence Intervals and Data Suppression

Confidence intervals have been provided for every statistic in the detail tables. These intervals indicate the margin for error because a sample was drawn rather than conducting a census. If the same general sampling procedures were repeated independently a large number of times, and a statistic of interest and its confidence interval were recalculated on each of those independent replications, the average of the replicated statistics would be contained within 95 percent of the calculated confidence intervals.

The confidence intervals reflected the effects of sampling and of the adjustments that were made to the weights. They did not generally reflect measurement variance in the questionnaires. The intervals were based on variance estimation techniques that will be available in separate technical reports. In brief, subsamples of the sample were drawn and put through the same estimation techniques. The adjusted variation among the subsamples provides an estimate of the variance of the total sample. Details on how confidence intervals were calculated from variance estimates follow.

Some estimates were suppressed. This was done when the reliability of a statistic was poor. This was measured in terms of the sample size and the width of the confidence interval. Estimated proportions near 0 percent and 100 percent were more likely to be suppressed than other estimates since it was difficult to estimate rare characteristics well. The exact criteria for this suppression also follow.

### A.3.1 Confidence Intervals

Variances were estimated for NSPY using a resampling approach. This resampling method was developed specially for NSPY. It uses 100 resamples to measure the variance in the full sample estimates. This method reflects, the variance due to selecting a larger sample of 100 PSUs for the standard Westat design, the variance due to subsampling to the 90 NSPY sample PSUs, and the variance due to sampling segment dwelling units and persons within PSUs. Moreover, it reflects the finite population correction factors at both the PSU and segment levels. Full technical documentation of this method can be obtained from Westat (Rizzo, 2000).

After each of the 100 resamples was drawn, the full set of adjustment procedures was run on each resample. This means that each resample was adjusted for nonresponse and was raked to adjusted

Current Population Survey (CPS) control totals. By doing this, the variance estimation procedure reflected the changes in uncertainty due to the point estimation procedures.

Once the variance estimates were obtained, they were translated into confidence intervals using approximations similar to those that have been developed on 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. The assumption of 100 degrees of freedom came from the 100 resamples. 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. Assuming 100 degrees of freedom was slightly more conservative. The standard error was multiplied by 1.98 instead of 1.96 to form a 95 percent confidence interval. The formula is

$$\text{lower bound} = \bar{x} - 1.98\sqrt{\text{var}(\bar{x})} \text{ and upper bound} = \bar{x} + 1.98\sqrt{\text{var}(\bar{x})}\,.$$

For proportions, it is assumed that a logistic transform of the estimated proportion has a normal distribution. This results in confidence limits that are strictly between 0 and 1, a useful property for estimated proportions. The formula for estimated proportions strictly between 0 and 1 is

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

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

For example, if the estimated proportion is 0.5 percent with a standard error of 0.4 percent, rather than calculating the standard t-approximation of -0.3 percent to +1.3 percent, the logistic formula yields a confidence interval of 0.1 percent to 2.4 percent.

Estimated proportions of 0 and 1 pose special difficulties for variance estimation and calculation of confidence intervals. The calculated variance estimate of zero is not meaningful for such estimated proportion, because the best confidence intervals are not collapsed at the point estimates. The approximation used for a confidence interval around an estimated zero proportion is

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

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), where n refers to the effective sample size defined to be the actual sample size divided by the average design effect (as suggested by D. Judkins and P. Zador). For these confidence intervals the average design effect was estimated to be 2.

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

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

As examples, if a domain has a sample size of 500, then the upper confidence limit on an estimate of 0 percent will be 1.5 percent and the lower confidence limit on an estimate of 100 percent will be 98.5 percent.

The Wave 3 confidence intervals for the counterfactual projections were done as in Wave 2. The youth counterfactual projections had standard errors estimated in WesVar using replication. However, the dyads counterfactual projections involved youth and parents so estimation for dyads was not as straightforward. The variance for youth and parents was estimated along with an estimate of the covariance between youth and parents based on replication. Once the estimate of standard error was obtained, the formula given above for the confidence intervals was used.

This report also contains confidence intervals for differences and means across waves. The samples in the three initial recruitment waves were independent except for PSU selection. For simplification, the PSU component of variance was ignored and the variances were assumed to be independent across waves. Both means and differences were approximated by assuming the statistic had a normal distribution and the t-distribution intervals with 100 degrees of freedom discussed above apply. For future waves, this will have to be changed because of the dependence of the statistics.

For differences of proportions where one or more of the estimates was 0 or 1 a slight modification of the above formula was needed. The approximation used for a confidence interval around an estimated zero proportion is

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

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 the only estimate was zero, the standard error for the nonzero estimate was used to impute the standard error for the zero estimate, adjusting for sample size.

## A.3.2 Suppression

There were several suppression criteria. All were developed with the aim of preventing overanalysis of statistics that contain little true information. For example, if a domain had a sample size of only two youth, and the estimated proportion of them who thought a certain way on some subject was 50 percent, then the confidence interval would range from 5.7 percent to 94.3 percent, which was too wide to be of any use.

Any estimate based on an effective sample size of 30 or less was suppressed. The effective sample size for a statistic was calculated as the simple random sample size of the same domain that would have generated a standard error of the same size.

Estimated proportions between 0 and .5 were suppressed if

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

and 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 these rules meant that larger effective sample sizes are required to avoid suppression as the estimated proportion approaches 0 or 1. Estimated proportions of 0 or 1 were suppressed if the effective sample size for the domain was 140 or less. This corresponds to confidence limits of (0.000-0.026) on 0 and (0.974-1.000) on 1.

## A.3.3 Average Design Effects and Effective Sample Sizes

A design effect is defined as the ratio of the achieved variance to the hypothetical variance that would have been achieved if a simple random sample of the same size had been used. An effective sample size is defined as the quotient of the nominal sample size divided by the design effect. Design effects were calculated for a number of statistics. They varied considerably from statistic to statistic, partially reflecting true differences in design effects but also reflecting substantial measurement noise. Table A-E shows the average design effects and corresponding effective sample sizes for statistics about youth, parents, and dyads.

**Table A-E**
**Design effects and effective sample sizes**

| Youth age domain | Youth | | Parents | | Dyads | |
|---|---|---|---|---|---|---|
| | Design effect | Effective sample size | Design effect | Effective sample size | Design effect | Effective sample size |
| 9-11 | 1.25 | 870 | 1.37 | 757 | 1.44 | 714 |
| 12-13 | 1.22 | 870 | 1.37 | 734 | 1.39 | 722 |
| 14-15 | 1.47 | 376 | Na | Na | 1.58 | 331 |
| 16-18 | 1.27 | 481 | Na | Na | 1.32 | 430 |
| 14-18 | 1.27 | 916 | 1.4 | 772 | 1.55 | 704 |
| Wave 1 Total | 1.46 | 2,268 | 1.66 | 1,882 | 2.27 | 1,374 |
| 9-11 | 1.27 | 727 | 1.38 | 634 | 1.38 | 626 |
| 12-13 | 1.26 | 522 | 1.28 | 483 | 1.31 | 469 |
| 14-15 | 1.49 | 264 | Na | Na | 1.49 | 250 |
| 16-18 | 1.46 | 265 | Na | Na | 1.58 | 227 |
| 14-18 | 1.49 | 524 | 1.50 | 484 | 1.69 | 443 |
| Wave 2 Total | 1.49 | 1,585 | 1.73 | 943 | 2.25 | 982 |
| 9-11 | 1.21 | 808 | 1.53 | 607 | 1.3 | 707 |
| 12-13 | 1.29 | 562 | 1.47 | 464 | 1.2 | 569 |
| 14-15 | 1.49 | 252 | Na | Na | 1.4 | 256 |
| 16-18 | 1.46 | 260 | Na | Na | 1.4 | 248 |
| 14-18 | 1.49 | 507 | 1.68 | 418 | 1.5 | 470 |
| Wave 3 Total | 1.64 | 1,499 | 1.82 | 923 | 2.0 | 1,153 |
| 9-11 | Na | Na | Na | Na | Na | Na |
| 12-13 | 1.18 | 636 | 1.62 | 384 | 1.35 | 473 |
| 14-15 | 1.21 | 759 | Na | Na | 1.87 | 406 |
| 16-18 | 1.29 | 550 | Na | Na | 1.95 | 282 |
| 14-18 | 1.43 | 1,309 | 1.46 | 784 | 2.24 | 584 |
| Wave 4 Total | 1.45 | 1,945 | 1.68 | 905 | 2.18 | 894 |

# References

(http://yawyl.claritas.com/).

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

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. (2000). Memorandum from Lou Rizzo to David Judkins. *NSPY variance estimation.*

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.

# Appendix B
# Data Collection Methodology and Response Rates

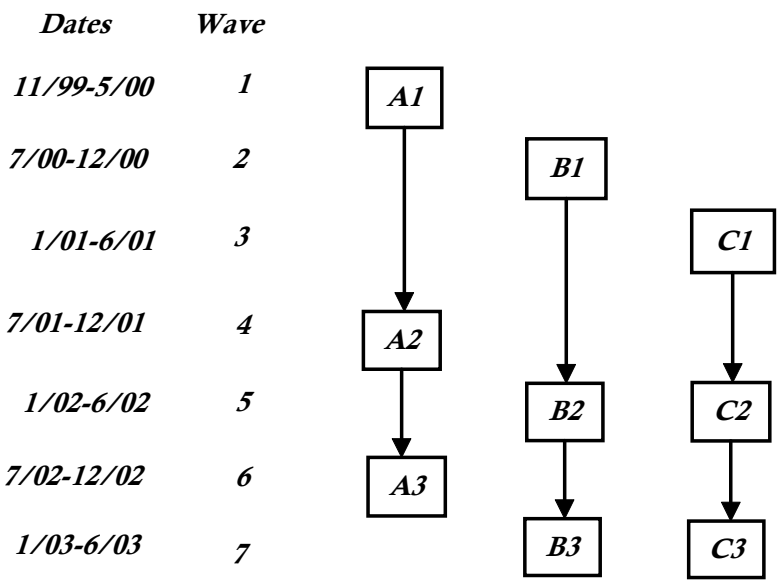
Two types of data were collected and analyzed for the Evaluation: quantitative survey data collected in a screener and three extended interviews (parent, teen, and child), and media buy data (i.e., Gross Rating Point (GRP) information).

This appendix describes the data collection methodology used during the three initial recruitment data collection waves (Waves 1 through 3) and the 1st followup wave (Wave 4) of the Evaluation. Topics include survey design, questionnaire design, pilot testing, interviewer recruitment and training, media activities, procedures used during data collection, data editing and cleaning, and response rates.

## B.1 Survey Design

The major evaluation component of the Phase III Evaluation is the conduct of the National Survey of Parents and Youth (NSPY), which is a longitudinal study, consisting of seven data collection waves, each lasting approximately six months. The NSPY is a nationally representative survey being conducted in 90 locations across the United States. Figure B-1 is a graphical depiction of the initial recruitment and followup plan of the NSPY.

**Figure B-1. NSPY INITIAL RECRUITMENT AND FOLLOWUP PLAN**



The initial recruitment phase (Waves 1 through 3) consisted of three cross-sectional surveys, lasting approximately 6 months each. During recruitment, approximately 81,000 households were screened for the presence of children in the age ranges of interest. Only about one in every eight households was determined was eligible to participate (12%).

The followup phase (Waves 4 through 7) began with the Wave 4 data collection. Parents and youth recruited during the first three recruitment waves are being tracked and recontacted two additional times during the followup. Wave 1 participants are being followed up in Waves 4 and 6. Wave 2 and Wave 3 participants are being followed up in Waves 5 and 7. The followup period can range from 6 to 24 months, depending on the wave and the dates of interview.

## B.2 Questionnaire Design

In preparation for the Evaluation of Phase III of the National Youth Anti-Drug Media Campaign, the National Institute on Drug Abuse (NIDA) convened an expert panel to assist in the development of data collection questionnaires. This group, which included specialists in adolescent drug use prevention and parenting behaviors, met and generated draft survey questionnaires for children (aged 9 to 11), teens (aged 12 to 18), and parents for the NSPY. NIDA shared these Phase III prototypes with Westat at the beginning of the contract period.

Westat formed a questionnaire development team whose members included evaluation experts from Westat, the Annenberg School for Communication at the University of Pennsylvania, and the National Development and Research Institutes (NDRI). This team reviewed the Phase III prototypes as well as the survey questionnaires used in the Phase II Media Campaign Evaluation, and other surveys, including Monitoring the Future (MTF), Community Action for Successful Youth, National Household Education Survey (NHES), and the National Household Survey on Drug Abuse (NHSDA).

To facilitate the development of the questionnaires, the team developed a behavioral change model for the Evaluation and mapped each question back to this model, as well as to the communication objectives that had been established for the Media Campaign.

Question domains for parents included the following:

- Media consumption;
- Past discussions with child about drug attitudes and avoidance strategies;
- Past child monitoring behaviors;
- Self-efficacy of discussing drugs with child and of monitoring the child's actions;
- Belief that the child is at risk for drug use;
- Belief that drug use has bad consequences;
- Exposure to the Media Campaign's advertising;
- Parent's own current and past use of tobacco, alcohol, and drugs; and
- Demographic information.

Youth question domains included the following:

- Exposure propensity to media;
- Youth's own current and past use of tobacco, alcohol, marijuana, and inhalants;
- Past discussions with and communication of anti-drug messages from parents and friends;
- Expectations of others about respondent's drug use;
- Knowledge and beliefs about the positive and negative consequences of drug use;
- Exposure to the Media Campaign's advertising;
- Family and peer factors;
- Personal factors; and
- Demographic information.

During Waves 1 and 2, virtually the same set of questions was asked of respondents. However, during Wave 3, some new questions were added. They included a question on brand recognition in the Teen and Parent questionnaires; questions about Ecstasy use in the Teen questionnaire (have used and when last used); questions about doing fun things with parents in the Teen and Child questionnaires; and a question about parents' perception of the efficacy of drug talk in the Parent questionnaire.

To make room for these questions, some questions were deleted. They included questions about reading magazines or seeing TV shows from the Teen and Parent questionnaires, questions about communicating rules for alcohol and smoking from the Teen and Child questionnaires, and a question about perceived consequences of inhalant use from the Child questionnaire.

In Wave 4, the extended interviews for parents, teens, and children were essentially the same as during Wave 3, except for some additional questions on Ecstasy in the Teen questionnaire. Additional Ecstasy questions included intentions to use, perceived expectations of use by peers, and attitudes of use including approval/disapproval of use and perceived harm of use.

The questionnaires for Waves 1 through 4 can be found on the NIDA web site: http://www.nida.nih.gov/DESPR/Westat/index.html. During Waves 1 through 3, a brief, hard copy household screening questionnaire was used to determine a sampled household's eligibility. All other data were collected using a laptop computer and a combination of computer-assisted interview technologies. Computer-assisted personal interview (CAPI) involved having the interviewer read the questions to the respondent and record the answers in the computer. In Waves 1 through 3, CAPI was used to enumerate the household and select a parent/caregiver and one or two youth. In Wave 4, CAPI was used to determine respondent eligibility and to select a new parent, if appropriate. CAPI was also used for the nonsensitive questions in the extended interview (parent, teen, and child) questionnaires in all waves. For collection of sensitive data in the extended interview questionnaires, audio computer-assisted self-interview (ACASI) technology was employed. This allowed respondents to self-administer the survey in total privacy. They listened to the question on headphones and recorded their own responses by touching the computer screen. These technologies were used based on the theory that providing respondents with a methodology that improved privacy and

confidentiality would make reporting of potentially embarrassing, stigmatizing, or illegal behaviors (such as drug use) less threatening, and enhance response validity and response rates.

On average in Waves 1 and 2, it took 6 minutes to enumerate and select household members for interview, and 34 minutes for children (aged 9 to 11), 44 minutes for teens (aged 12 to 18), and 52 minutes for parents to complete their respective extended survey questionnaires. The above noted changes to the Wave 3 questionnaires resulted in the following timings: 6 minutes to enumerate and select household members; and approximately 31 minutes for children, 41 minutes for teens, and 55 minutes for parents to complete their respective extended questionnaires. In Wave 4, it took on average 6 minutes to complete the computerized screener. Timings for the extended instruments were approximately 35 minutes for children, 44 minutes for teens, and 54 minutes for parents.

## B.3 Pilot Test

Once the Office of Management and Budget (OMB) clearance was obtained, Westat conducted a pilot test in Baltimore, Maryland, prior to Wave 1. Approximately 300 households were screened to obtain about 20 household interviews using the NSPY questionnaires. The purpose of the pilot was to test the adequacy of questionnaire skip patterns, question wording and flow, and test the application of the ACASI portion of the questionnaire as well as the adequacy of the advance materials and interviewing procedures. A debriefing was held at the end of the pilot data collection. From that, some questions needed to be dropped from each of the extended interview questionnaires to keep within the OMB respondent burden estimates. Procedures and advance materials were updated as appropriate.

Westat conducted a second pilot test prior to Wave 4 to test its Followup screening instruments. The participating households in the first Baltimore pilot test were recontacted and screened for Followup status. The purpose of the pilot was to test the adequacy of the screening instruments, skip patterns, question wordings and flow, as well as advance materials and interviewing procedures. An interviewer debriefing was held at the end of this pilot data collection as well. From that debriefing, some minor changes were made in Followup screening questions and procedures.

## B.4 Interviewer Recruitment and Training

The NSPY initial data collection design was based on hiring one primary interviewer in each of 90 primary sampling units (PSUs) and hiring approximately 35 more interviewers to supplement efforts in larger PSUs, PSUs geographically clustered, and in PSUs where primary interviewers quit during the field period. Twenty-nine additional interviewers were hired to supplement the data collection effort later in Wave 1. No additional interviewers were needed to staff Wave 2. Subsequent interviewer attrition required that 26 additional interviewers be hired to supplement the data collection effort in Wave 3. In Wave 4, 28 interviewers were fielded at the beginning of the Wave and three more were hired and trained during the wave.

Initially, interviewers were recruited from Westat's pool of experienced interviewers. Additional candidates were recruited through local organizations and classified newspaper advertisements placed in various PSUs as needed. These candidates were screened for communications skills and availability. Spanish language interviewer candidates were screened by bilingual project staff for their ability to communicate effectively in both Spanish and English. Approximately 12 percent of the total

interviewers hired were bilingual. Most English and bilingual candidates had prior experience relevant for data collection.

Over the waves, all interviewers participated in an 8 to 10 day training session. The training program, which was staffed by qualified project staff and field supervisors, was designed to ensure consistency in data collection through the use of lectures, with a heavy focus on practice sessions. Trainees new to Westat attended an additional half-day training on general interview techniques. Bilingual trainees also attended an additional half-day training that concentrated on reviewing bilingual scripts and materials.

## B.5 Media Activities

Because this is an evaluation of a media campaign, activities such as media buying, ad creation, and broadcast levels play key roles in the questionnaires as well. Because the Media Campaign is dynamic over time, the media-specific questions in the questionnaires must also change appropriately.

In the evaluation of Child, Teen, and Parent questionnaires, some questions are asked about the respondent's media usage patterns, including television, radio, and magazines. All NSPY questionnaires contain a section of questions devoted to how the respondent receives anti-drug messages. In these questions, selected television and radio Media Campaign ads that have been broadcast during the prior 2 calendar months are played for the respondent. Questions are then asked about the respondent's recall of prior exposure (viewing or listening) to the ad, and his/her assessment of the ad's message and impact. The set of television and radio ads that are played for respondents are changed monthly, with a set protocol being used to determine which ads are played during each month and for which respondents.

Each month Ogilvy, the Campaign media buy contractor, produces an updated copy rotation schedule. This schedule outlines, by month, each ad that is slated for broadcast, its target audience (parents or youth), and racial or ethnic group (general market, African American, or Hispanic). Included are each ad's planned broadcast dates and the Media Campaign behavioral platform that the ad addresses. As ads are produced, Ogilvy forwards them to Westat for digitizing; a process that puts the ads into an electronic format that can then be incorporated into the computerized laptop questionnaires.

Using the current copy rotation schedule, Westat determines those television and radio ads that will need to be played to respondents over the next 2 months. A CD containing those ads is then produced and sent to the field interviewing staff. A look-up table is also developed for each interview month and transmitted to the field staff. It provides the specifications for ad selection and randomization for each respondent that month.

During Waves 1 and 2, questions were asked about viewership of specific television shows and readership of specific magazines from which Ogilvy purchased advertising time or space. The specifics of these media buys were determined based on the Gross Rating Points (GRPs) that the television shows, radio program or magazine were expected to earn. Ogilvy sent updated information on those television shows and magazines for which ad time or space has been purchased to Westat every 3 months, and appropriate updates were transmitted to the field interviewers' laptop questionnaires. (Questions on specific television shows and specific magazine readership were dropped from the Wave 3 and Wave 4 questionnaires.)

Ogilvy also provides data regarding the planned GRP levels for the previous quarter, by target audience (parents or youth), creative ad execution, media (television, radio, print, and out of home), and week/month. GRPs refer to the percentage of the target population that is estimated to be watching a particular TV show, listening to a specific radio program, or reading a certain magazine, and are therefore exposed to the advertising messages provided. These GRPs are based on data from that media's audience ratings company (Nielsen Media Research for television, Arbitron Research and RADAR for radio and MRI for print). Knowing the reach and frequency objectives for the Media Campaign's messages, the media buyers then purchase a mix of media whose GRPs, when aggregated, should achieve the desired intensity of Media Campaign message exposure. This information is used by the Evaluation's analysts to look for correlation between recalled exposure to ads by respondents and the ads' reach and frequency levels.

## B.6 Initial and Followup Data Collection

This section discusses five topics central to the NSPY data collection effort. They include the procedure used to determine whether the household was eligible to participate in the survey, the rules adopted for collection of information from neighbors pertaining to household eligibility determination, how household members were subsampled for inclusion as survey respondents, steps taken to assure respondent confidentiality, and the procedures used to validate Waves 1 through 4 interviews.

### B.6.1 Determining Household/Respondent Eligibility

During Waves 1 through 3, interviewers were required to make up to five in-person attempts to contact a household. A household was considered eligible if two criteria were met. First, the household must contain children of a specified age group (age groups included households with children aged 9 through 13, 12 and 13, or 9 through 18). Second, the housing unit must have been built before April 1, 1990 be a mobile home, or be selected through the permit sample (see Appendix A). All eligibility information was collected hard copy and then entered into an electronic file on laptop computers.

To be included in the Wave 4 Followup sample, a household must have had at least one selected person (parent, teen or child) complete his or her extended interview in Wave 1. If no one who was selected completed an interview in Wave 1, then the household was not included in the Wave 4 Followup sample.

Prior to Wave 4, efforts were made to verify the location of Wave 1 adult respondents. Location information (i.e., address and telephone number) about Wave 1 respondents was sent to a National database company for tracking purposes. Updated location information from this source was sent to Westat's Telephone Research Center and telephone interviewers placed calls to these households to verify the identity of the Wave 1 respondents.

At Followup, interviewers were allowed to screen households both by telephone and in-person. Interviewers were required to make up to five telephone attempts to contact a household. If the telephone attempts were not successful, up to 5 in-person attempts were then made. Most first attempts were made by telephone, however first attempts at contact were made in-person if the selected parent had refused to complete his or her initial interview or if the interviewer did not have a telephone number to call.

A youth who had been selected at Wave 1 was considered eligible for the Wave 4 survey if the youth was 9- to 18 -years of age at the time of the Wave 4 interview and was not living in a group quarters situation (that is was not living away from home at school or in an institution). A parent or caregiver who had been selected at Wave 1 was considered eligible for the Wave 4 survey if he or she was still living with an eligible sampled youth at least two nights a week and was not physically or mentally disabled. A new parent was chosen for interview if either of these two conditions was not met.

## B.6.2 Use of Neighbor Reports to Determine Eligibility

Through most of the initial data collection waves, interviewers were instructed to visit the sampled household three times to try to determine eligibility, prior to obtaining eligibility information from a neighbor. This procedure was changed for a short period of time during Waves 1 through 3 to allow interviewers to determine eligibility information from neighbors after one attempt to contact the household. Because a neighbor might be less able to accurately know the exact ages of children, two questions about children were asked. First, the neighbor was asked whether any children aged 9 to 18 lived in the household. If yes, a followup question was asked to determine whether children of the specified age for the particular household (see categories above) lived in the household. In addition, the neighbor was asked if sampled housing units in area segments were built after April 1, 1990. Finally the neighbor was asked what times members of the sampled household would be likely to be at home. If answers to both of the age questions were no, the household was considered ineligible. If the answer to either or both age questions was yes and if the housing unit was built before April 1, 1990, or if the housing unit was drawn from the permit sample, the interviewer continued to try to contact the sampled household. Remaining attempts were made to contact the sampled household to obtain an interview at times suggested by the neighbor.

Neighbor reports to determine eligibility were not applicable to the followup data collection waves.

## B.6.3 Selection of Respondents

During Waves 1 through 3, the interviewer conducted a household enumeration with a household member 18 years of age or older, once a household was determined to be eligible. All members of the household, excluding children/students who were currently away from home, living at a boarding school or college, were enumerated. At this point, up to two eligible children were randomly selected. Once the children were selected, the relationship of every other person to the selected child was obtained. One or two parents or primary caregivers were then selected based on a predetermined algorithm. (Two parents or primary caregivers were chosen only in the unusual situation where the selected youth were not siblings.) If two parents for a selected child resided in the household, the algorithm selected the male or female parent on a random basis. If one of the parents was a stepparent or foster parent, that parent must have lived with the child in the household for a least 6 months to be eligible for selection. If no parents lived in the household, the algorithm selected a primary caregiver. Once all respondents were selected, information on the race and ethnicity for each selected person was obtained.

As mentioned earlier, youth were considered eligible for the Wave 4 survey if they were 9- to 18- years old at the time of the Wave 4 survey and were not living in group quarters. New youth were never selected as replacements for ineligible ones. New parents/caregivers, however, could be selected in Wave 4, if the Wave 1 parent/caregiver was ineligible for the survey at Wave 4.

For all waves, all respondent selection information was entered into a laptop by the interviewer using a CAPI approach.

## B.6.4 Guaranteeing Confidentiality

An important part of the survey methodology was to obtain honest answers to very sensitive data. To meet this end, several procedures were implemented. First, a Certificate of Confidentiality was obtained for the study. Under the certificate, the Federal Government pledged that the evaluation team cannot be compelled by any person or court of law to release a respondent's name or to link a respondent's name with any answers he/she gives. Interviewers showed a copy of the certificate to respondents prior to the interview. They also guaranteed that all respondent names and other identifying information would be destroyed at the end of the study and would not appear in any publications resulting from the study. Teen and child assent forms were appropriately worded for each age group to make sure that the youth understood that the answers they gave would be kept private and would not be connected with their names.

Second, the extended interviews were administered in a CAPI and ACASI format. Sensitive questions were in ACASI format, which meant that respondents used the computer themselves to answer questions by touching the screen and used headphones to hear the questions. The extended interview was programmed so that the interviewer was unable to go back into the interview and look at answers the respondent provided in the ACASI section.

Third, interviewers were instructed to, if possible, seat the respondent in a chair that was against the wall or a piece of furniture so that no other person could stand or pass behind the respondent. This procedure hindered third parties from being able to observe the respondent's answers during the ACASI part of the interview. The interviewer also requested that parents not be present in the room while the questionnaire was being conducted with the youth. If the parent insisted on being present in the room, the interviewer asked the parent not to stand directly behind the child during the ACASI portion of the interview.

## B.6.5 Validation of Interviews

During Wave 1, 10 percent of parents interviewed were selected for validation. Approximately 75 percent were contacted by telephone and attempts to contact the remainder were made by mail. When interviewers were suspected of falsifying data, all of their worked cases were redone by different interviewers. In a few instances, interviewers were terminated for falsifying data.

During Wave 2, approximately 13 percent of parents interviewed and 2 percent of the ineligible households were selected for validation. Approximately 58 percent were contacted by telephone, and attempts to contact the remainder were made by mail. No invalid cases were found during Wave 2.

During Wave 3, approximately 18 percent of the parents interviewed and 5 percent of the ineligible households were selected for validation. Approximately 76 percent were contacted by telephone and attempts to contact the remainder were made by mail. When an interviewer was suspected of falsifying data, all of his or her worked cases were redone by different interviewers. In one instance, an interviewer was terminated for falsifying data.

During Wave 4, approximately 13 percent of the parents interviewed and 44 percent of the ineligible households were selected for validation. Approximately 86 percent were contacted by telephone and attempts to contact the remainder were made by mail. No invalid cases were found for interviewers completing Wave 4 work, however two interviews completed during Wave 1 were identified as questionable during Wave 4 when an interviewer revisited the households.

## B.7 Data Editing and Cleaning

SAS programs were developed to perform edit checks on the screener and extended interview data. All interview skip patterns were checked to ensure that data did not exist for data items that should have been skipped and that data values were missing only when a data item had been properly skipped. Checks were also performed to confirm that all reported ages and dates were in a logical sequence between birth and the date of interview. Additional edit checks were executed to ensure that questions were asked regarding the appropriate groups of ads, given the demographic characteristics of the respondent. After the SAS edits were reviewed and the appropriate updates were applied, frequencies were produced for all variables at the dwelling unit level, the sampled person level, and the parent/youth dyad level. These frequencies were reviewed by experienced data specialists who identified outliers, unexpected missing data, and data inconsistencies. When a potential problem was identified, the data manager located the corresponding records within the database and evaluated the data to determine if any items needed to be updated.

Data updates were recorded by the data specialists and were carried out through a SAS update program that updated the appropriated data items and kept a transaction record of all updates.

## B.8 Response Rates

### Wave 1

There were 34,691 sampled addresses to be contacted and screened in NSPY Wave 1. Of those sampled addresses, 4,649 (13.4%) were discovered to be either vacant or nonresidences (such as businesses or other institutions). That left 30,042 occupied residential addresses to be contacted and screened for study eligibility.

Of those occupied addresses, answers to the screening questions were obtained for 28,567 (95.1%). Roughly 1 in 8 screened addresses (12.2%) had children in the required age ranges and were eligible to participate in NSPY.

In the 3,497 eligible households, data collection staff were able to enumerate household members for 2,602 (74.4%) households, so that a parent/caregiver and one or more youth could be selected for interview. Once selected 2,293 (88.4%) of NSPY parents/caregivers completed an interview. Interviews were completed with 3,312 (90.6%) of selected NSPY children and teens.

The cumulative response rate (screener response rate x roster response rate x interview response rate) was 64.1 percent for youth and 62.5 percent for parents.

## Wave 2

There were 23,000 sampled addresses to be contacted and screened in NSPY Wave 2. Of those sampled addresses, 2,405 (10.5%) were discovered to be either vacant or nonresidences (such as businesses or other institutions). That left 20,595 occupied residential addresses to be contacted and screened for study eligibility.

Of those occupied addresses, answers to the eligibility screening questions were obtained for 19,701 (95.7%). Roughly 1 in 8 screened addresses (12.7%) had children in the required age ranges and were eligible to participate in NSPY.

In the 2,502 eligible households, data collection staff were able to enumerate household members for 1,866 (74.6%) households, so that a parent/caregiver and one or more youth could be selected for interview. Once selected, 1,632 (88.2%) of NSPY parents/caregivers completed an interview. Interviews were completed with 2,362 (91.9%) of selected NSPY children and teens.

The cumulative response rate (screener response rate x roster response rate x interview response rate) was 65.6 percent for youth and 62.9 percent for parents.

## Wave 3

There were 23,300 sampled addresses to be contacted and screened in NSPY Wave 3. Of those sampled addresses, 2,272 (9.8%) were discovered to be either vacant or nonresidences (such as businesses or other institutions). That left 21,028 occupied residential addresses to be contacted and screened for study eligibility.

Of those occupied addresses, answers to the screening questions were obtained for 20,085 (95.5%). Roughly 1 in 8 screened addresses (12.8%) had children in the required age ranges and were eligible to participate in NSPY.

In the 2,566 eligible households, data collection staff were able to enumerate household members for 1,931 (75.3%) households, so that a parent/caregiver and one or more youth could be selected for interview. Once selected, 1,681 (87.6%) of NSPY parents/caregivers completed an interview. Interviews were completed with 2,459 (91.2%) of selected NSPY children and teens.

The cumulative response rate (screener response rate x roster response rate x interview response rate) was 65.5 percent for youth and 63 percent for parents.

## Wave 4

Four separate response rates were calculated for Wave 4. These include:

- A followup cross-sectional response rate;
- A cumulative cross-sectional response rate;
- A followup longitudinal response rate; and
- A cumulative longitudinal response rate.

Under the NSPY sample design, subsamples of youth and parents selected during Wave 1 were retained for followup in Wave 4. For the cross-sectional survey, youth and parents in households that completed a screener roster in Wave 1 were included in the followup sample if the household contained at least one Wave 1 sample person (either parent or youth) who completed an interview. As a result, under the selection criterion employed for Wave 4, a small number of youth and parents sampled at Wave 1 who did not complete a Wave 1 interview were refielded in Wave 4. These "extra" youth and parents were used only for the cross-sectional analysis and, therefore, were accounted for in the cross-sectional response rate. For the longitudinal analysis, a youth and parent must have completed an interview in Wave 1 and in Wave 4 to be included in the calculation of the longitudinal response rate.

## Cross-Sectional Response Rates

### Followup Cross-Sectional Response Rate (FCRR)

The FCRR represents the percentage of parents and youth that were successfully located and interviewed during Wave 4 of the sample fielded in Wave 4. It is defined as:

$$\text{FCRR} = \frac{\text{\# Households Completing Eligibility Screening}}{\text{\# Households Fielded}} \times \frac{\text{\#Respondents Completing Interview}}{\text{\# Respondents Eligible to Participate}}$$

There were 2,602 households that completed the household enumeration (roster) screening at Wave 1. Based on data collected during Wave 1, 2,450 (94.2%) of these households contained at least one respondent from Wave 1 (either a youth or a parent) and thus were eligible for refielding at Wave 4. The further exclusion of households that contained only youth who were expected to be age 19 or older at the beginning of the Wave 4 data collection resulted in the refielding of 2,304 households in Wave 4.

Followup telephone or inperson eligibility screening was attempted for the 2,304 households that were refielded in Wave 4. Of these, eligibility was determined for 1,999 (86.8%) of the households. For the remaining 305 households, eligibility could not be determined for various reasons (e.g., the household moved out of the interviewing area or was not locatable, the household could not be contacted for some other reason, or the household refused to complete the eligibility screener.)

The 1,999 successfully screened households contained 2,744 Wave 1 youth, of which 96 (3.5%) youth were determined to be ineligible for the Wave 4 survey (e.g., were 19 years or older, were institutionalized or living in group quarters, or were deceased). Of the 2,648 eligible youth in the screened households, 2,478 (93.6%) completed the Wave 4 interview. Corresponding to the 2,648 youth, 1,939 parents were identified and 1,752 (90.4%) of them completed the Wave 4 interview.

Thus, the followup cross-sectional response rate for Wave 4 youth is **81.2** percent (86.8% x 93.6%); and the followup cross-sectional response rate for Wave 4 parents is **78.5** percent (86.8% x 90.4%).

### Cumulative Cross-Sectional Response Rate (CCRR)

The CCRR is the combination of the Wave 1 and Wave 4 survey response rates. It is defined as the product of the five following rates:

- The percentage of households at Wave 1 where eligibility was determined.

- The percentage of eligible households at Wave 1 where the household roster was completed;
- The percentage of Wave 1 households that were refielded (i.e., contained at least one respondent at Wave 1) at Wave 4;
- The percentage households at Wave 4 where eligibility screening was determined; and
- The percentage of youth/parents who completed the Wave 4 interview.

Thus, the cumulative cross-sectional response rate for Wave 4 is **54.1** percent (95.1% x 74.4% x 94.2% x 86.8% x 93.6%) for youth and **52.2** percent (95.1% x 74.4% x 94.2% x 86.8% x 90.4%) for parents.

## Longitudinal Response Rates

### Followup Longitudinal Response Rate (FLRR)

The FLRR represents the percentage of parents and youth that were successfully located and interviewed in Wave 4, who were also successfully interviewed in Wave 1. It is defined as:

$$\text{FLRR} = \frac{\text{\# Respondents where Eligibility Determined}}{\text{\# Respondents Interviewed in Wave 1}} \times \frac{\text{\# Respondents Completing Interview}}{\text{\# Respondents Eligible to Participate}}.$$

Of the 3,072 youth completing the Wave 1 who were refielded in Wave 4, eligibility status was determined for 2,685 (87.4%) youth. Of those youth, 96 were determined during Wave 4 screening to be ineligible for the Wave 4 survey (e.g., were 19 years or older, were institutionalized or living in group quarters, or were deceased). Among the 2,589 eligible youth, 2,435 (94.1%) completed the Wave 4 interview. Similarly, of the 2,158 parents completing the Wave 1 interview that were refielded in Wave 4, eligibility status was determined for 1,885 (87.3%) parents. Of those parents, 93 were determined during screening to be ineligible for the Wave 4 survey. Among the 1,792 eligible parents, 1,644 (91.7%) completed the Wave 4 questionnaire.

Thus, the followup longitudinal response rate for Wave 4 youth is **82.2** percent (87.4% x 94.1%); and the followup longitudinal response rate for Wave 4 parents is **80.1** percent (87.3% x 91.7%).

### Cumulative Longitudinal Response Rate (CLRR)

The CLRR is the combination of the Wave 1 and Wave 4 response rates based on a subset of respondents, i.e., those respondents who were interviewed in both Wave 1 and Wave 4. It is defined as the product of the three following rates:

- The cumulative Wave 1 response rate;
- The percentage of youth/parents at Wave 4 for whom eligibility was determined; and
- The percentage of eligible youth/parents who completed the Wave 4 interview.

Thus the cumulative longitudinal response rate for Wave 4 is **52.7** percent (64.1% x 87.4% x 94.1%) for youth and **50.1** percent (62.5% x 87.3% x 91.7%) for parents.

# Appendix C
# Methodology for Confounder Control

## C.1 Introduction

In this report, there has been considerable focus on changes in exposure and outcomes over time. If positive change occurs, then one wonders what might have led to the change. The level of exposure informs us about the activity level of the Campaign. It becomes more plausible to attribute some of the credit for any positive changes in outcomes to the Campaign if high exposure levels are attained and sustained. Most importantly, if people with higher exposure doses have better responses, it becomes plausible to believe that the treatment caused the response to be different from what it would have been in the absence of the Campaign. In the case when exposure and outcomes are measured simultaneously, the method provides important support for an inference of Campaign effect if one can assume that no other variable accounts for the observed association of exposure and outcome, and that the association is not the result of the outcome causing the exposure rather than vice-versa. This type of analysis is sometimes called a study of the dose-response relationship, analogous to a drug study comparing a 40 mg dose to a 20 mg dose.

Section C.2 discusses the strengths and weaknesses of the dose-response approach. Section C.3 provides more detailed information about the procedures used to implement it. Section C.4 provides detailed technical information on how effects were estimated. Section C.5 provides detailed technical information on how confidence intervals were formed on the effect estimates and how hypothesis testing was conducted.

## C.2 Strengths and Weaknesses of the Dose-Response Approach

Interpretation of change over time in outcomes relies on the assumption that other factors (everything other than the Campaign) affecting drug-related cognitions and use held steady during the time period. However, it was beyond the scope of this evaluation to determine whether forces external to the Campaign did hold steady. These external forces might include such things as drug prices, drug availability, content of popular media, content of political speech and debate, celebrity actions, and seasonal variations. Consequently, the required assumption of constancy in all other societal forces is a strong assumption. Furthermore, data collection started after the start of the national phase (Phase II) of the Campaign. So even if one were to accept the strong assumption about other forces holding steady, change in outcomes would reflect only the incremental effect of additional exposure beyond any effect that could have been initially achieved. Given these caveats, it is clear a positive trend, while desirable, is insufficient for evaluating the effectiveness of the Campaign. Similarly, a negative trend does not negate the possibility that campaign effects existed, but countervailing effects from other causes were stronger.

In this report, we discuss trends over time but the principal analytic approach taken was to study the dose-response relationship, where the dose is a unit of exposure to anti-drug advertising, and the response is the simultaneously observed cognitive variables about drug use or parenting practices. This approach is common in the epidemiology of chronic conditions brought on by environmental factors such as coal dust, primary smoking, second-hand smoke, indoor radon gas, and so on. The underlying theory in those disciplines is that if a substance is toxic, then a large dose of it should be at least as toxic as a small dose. If this expected relationship does not hold, the toxicity of the material has not been demonstrated. In the application of this theory to our evaluation of the Media Campaign, the underlying theory is that if advertising is effective, a large dose of consumed advertising should be at least as effective as a small dose. If this relationship does not hold, then Evaluation generally cannot conclude that the effectiveness of the advertising has been demonstrated.

In dose-response analysis, one must assume that the variation in doses is random after controlling for known factors. In randomized experiments such as clinical trials, random assignment within groups of substantive interest is used to ensure that doses are randomly given. However, since Media Campaign doses are not randomly assigned, but are instead self-chosen by choices in media consumption and filtered through subject's recall, the Evaluation must instead assume that all sources of systematic (nonrandom) variation in doses have been measured.

This is a strong assumption, but as part of the questionnaire design and acquisition of geographic information, the Evaluation team considered a wide range of background variables that might affect dose reception. However, there is always the risk that the questionnaires might not have measured all the predisposing variables. The questionnaires for Waves 1 through 3 can be found on the NIDA web site: http://www.nida.nih.gov/DESPR/Westat/index.html. Researchers can scan the list of questions that were asked and think about what might have been left out. Leaving important predisposing variables out of the analysis means that false effects can emerge from the dose-response study. The Evaluation team tried to include as many variables as seemed to be plausible predisposing variables, but limitations on the length of each interview meant information could not be recorded about every plausible predisposing variable.

Even among the set of data collected, some of the data items were not allowed into the "pool of admissible predisposing factors." This was necessary because some of the variables that were measured had an unclear temporal order with the outcomes. Some may be consequences of exposure to Campaign messages. Controlling on such "mediating" variables would be to underestimate Campaign effects. For example, if watching Campaign ads leads youth to change their beliefs about the consequences of marijuana use, and these belief changes lead, in turn, to changes in intentions to abstain from marijuana use (as would occur under the theoretical model described in Chapter 2), then it would be a serious mistake to allow marijuana beliefs into the pool of admissible predisposing factors, even though it is true that beliefs are predisposing factors in developing intentions about marijuana abstention.

Because the data for the first three waves were collected in a single session with each respondent, the internal causal ordering of data was often ambiguous. At this point in the process, human judgment was required to decide which variables were potential mediating variables and which were predisposing variables that were not subject to influence by exposure to the Campaign. There were some variables for which valid arguments were advanced both for classification as a mediator and for classification only as a confounder. Resolving such conflicts was difficult and of the utmost importance, because each decision potentially affects the evaluation findings. The Evaluation team

recognized that other researchers may disagree with these choices. A few of the decisions were extraordinarily difficult to make and are discussed in detail below.

At this point, it was decided as a matter of evaluative protocol not to experiment with alternate decisions. The decisions made for the second semiannual report are still those used for this fourth semi-annual report. Section C.3 presents an expanded discussion of each decision.

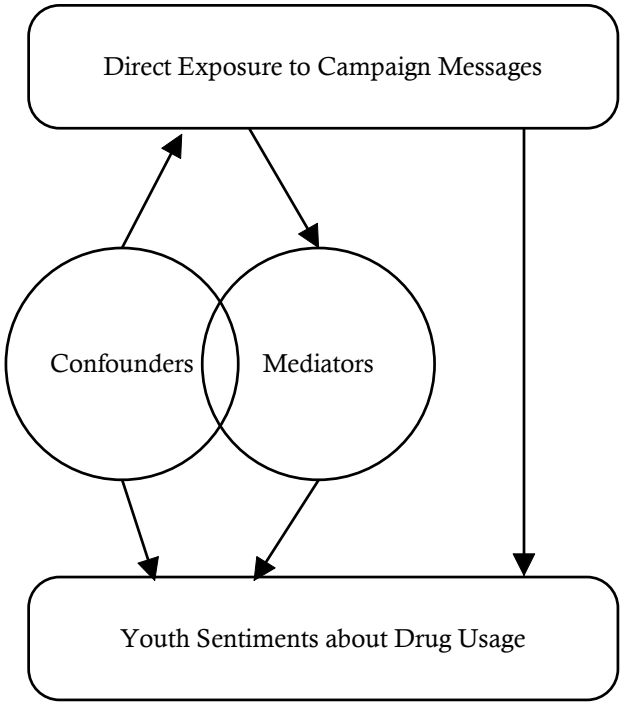
# C.3 Admissible Confounder Selection

This section presents the set of variables that the evaluation team admitted into analysis as youth confounders, the set accepted as parent confounders, and concludes with a list of confounders considered as potential moderators. The presentation commences with a brief discussion of the concept of confounding and moderating variables and of the analytic difficulties that arise because some variables may play both roles.

## C.3.1 Confounders and Mediators

A large number of cognitive and behavioral variables were obtained on each subject at a single point in time. It is impossible to say with any certainty the order in which these cognitions and behaviors manifested themselves in each subject. Nonetheless, in order to make causal inferences, it is necessary to make some assumptions about this ordering. Figure C-A defines different types of variables schematically.

**Figure C-A. Types of variables**




A confounder is a variable that leads to variation both in exposure and in outcomes but is itself not caused by exposure or outcomes. This is illustrated in Figure C-A by the directions of the line—

confounders cause variation in exposure and cause variation in anti-drug sentiments. In order to avoid false claims of Campaign effects as well as false claims of counterproductive Campaign effects, it is essential to remove the (confounding) effects of the confounder from the study of the dose-response relationship. Examples are given in the prior section of how this works.

A mediating variable is one that is associated with both exposure and an outcome, as is the case with a confounding variable, but a mediating variable is a result of exposure rather than a cause of exposure. This is illustrated in Figure C-A by the direction of the arrow connecting Mediators and Exposure. In other words, the mediating variable is causally posterior to exposure rather than causally prior to exposure. In order to prevent errors of omission where we do not identify a Campaign effect, it is vitally important that nothing be done to remove the (mediating) effects of the mediator from the study of the dose-response relationship.

Unfortunately, some variables play both confounding and mediating roles. This is illustrated in Figure C-A by the overlap of the circles for confounders and mediators. For variables in this overlap area, we have conflicting imperatives. We must both remove and not remove their effects. As an example of a variable in that overlap, consider the role of cigarette smoking. Cigarette smoking makes it easier to try marijuana and could be related to choices of TV and radio programs and hours of viewing—so it is a confounding variable. At the same time, there may be kids who stopped smoking or were prevented from smoking because of generalized effects of exposure to the Campaign as discussed in Chapter 2.

Thus, it is also a mediating variable. We included items as confounders only when we could be confident that they were not mediators. In the case of cigarette smoking, the issue was resolved by including smoking initiation if it occurred more than 1 year before the date of the interview.

Decisions about which variables would be regarded as potential confounders and which as mediating were made after discussion by a committee of the evaluation team prior to any examination of the data. The committee did not use any of the data about the relationships among the potential confounders/mediators, exposure, and outcomes in making these decisions. Thus the decisions were made blinded to any possible effects on either finding or not finding any effects of the Media Campaign.

## C.3.2 Admissible Pool of Youth Confounders

The following variables were judged by the committee to properly belong in the pool of admissible potential confounders for youth. The included variables can be divided into two broad groups. The first group, listed immediately below, include confounders that directly measure the respondent youth's personal demographics, attitudes, family environment, and behaviors. Discussion of particular exclusion and inclusion decisions follow the list.

1. Age
1. Gender
1. Race ethnicity
1. Neighborhood characteristics from the census
1. Urban, suburban, or rural nature of neighborhood
1. School enrollment status in the previous year

1. Whether school was in session in the last 30 days
1. Number of missed schooldays due to illness in the previous 30 days
1. Number of days the youth cut school in the previous 30 days
1. School grade level
1. Academic performance
1. Participation in extra-curricular activities[1]
1. Respondent's primary post-secondary plan
1. Hours of TV consumption on weekdays
1. Hours of TV consumption on weekends
1. Hours of radio consumption on weekdays
1. Hours of radio consumption on weekends
1. Internet use
1. Magazine reading habits
1. Language of TV viewing
1. Language of radio programs heard
1. Availability of cable or satellite TV in the household
1. Consumption of specific cable channels targeted by the Media Campaign
1. Personal assessment of family fighting
1. Personal assessment of feelings of family togetherness
1. Degree of parental supervision
1. Respondent's perception of parental knowledge of his or her activities
1. Respondent's perception of parental knowledge of his or her plans
1. Degree of enjoyment of time spent with his or her family
1. Youth rating of the importance of religion in their lives
1. Attendance of religious services
1. Personal antisocial behavior
1. Association with antisocial peers
1. Youth close friends' drug use
1. Personal tobacco use of a long-standing nature
1. Personal alcohol use of a long-standing nature
1. Sensation seeking tendencies.

[1] It has recently been argued that some of the Campaign advertising in early 2001 may have encouraged youth to join extra-curricular activities and thus, that this variable should be treated as a mediator rather than a confounder. This seemed of much less plausibility than a concern that such activities might both affect access to advertising as well as patterns of drug beliefs and use. The committee assumed that participation in extracurricular activities was largely a function of opportunity, physical fitness, other personal traits, accidents of friendship, and parental memories about extracurricular activities.

All of the above reflect youth reports about themselves, their friends, and their families. Some of these variables might be possible outcomes of drug use, and it could be argued that if the Campaign had reduced drug use these were posterior to the Campaign not prior to it. However, all of the analyses reported in Chapter 5 focus on youth who had not yet used drugs, thus the concern is reduced. For example, the Campaign might potentially reduce drug use and that might decrease family tension and increase a feeling of togetherness. Controlling for family togetherness might reduce that apparent dose response effect of the Campaign. However, given that only nonusing youth are studied in Chapter 5, family togetherness is appropriately seen as a confounder. Still some of these variables, contrary to the Evaluation team's considered judgment, might be causally posterior to either exposure or outcomes and thus not be true confounding variables.

Youth attendance of anti-drug programs (in or out of school) were excluded as confounders. There was some risk that youth reports of attendance at such programs might reflect access to Campaign advertising or other outreach efforts, particularly since substantial advertising buys were made on Channel One, an in-school network.

The second broad category included as admissible potential confounders for the youth analysis covers information on parental characteristics and perceptions. These included:

1. Parental age
1. Parental gender
1. Parental marital status
1. Parent has a child aged 9 to 11
1. Parent has a child aged 12 to 13
1. Parent has a child aged 14 to 18
1. Parental income
1. Parental educational attainment
1. Parental religiosity
1. Sharing of parental responsibilities
1. Parental use of the Internet
1. Parental consumption of newspapers
1. Parental consumption of magazines
1. Parental consumption of TV
1. Parental radio consumption
1. Parental consumption of specific cable channels targeted by the Media Campaign
1. The primary language in which the parent watches TV
1. Parental assessment of family togetherness
1. Parental enjoyment of time spent with children
1. Parent's perception of fights with children
1. Parent-child participation in fun indoor activities
1. Parent-child participation in fun outdoor activities

1. Parent's reports on the respondent youth's grade level
1. Parent's report on child's academic performance
1. Parent's report on the time their child spends with friends
1. Parental alcohol use
1. Parental tobacco use
1. Parental prior or current use of hard drugs
1. Parental prior or current use of marijuana
1. Parental prior or current use of inhalants

As with the youth variables, some of these variables have an ambiguous causal order with respect to outcomes and exposure. The fact that all the youth in the associational analysis are nonusers of marijuana strongly mitigates these concerns, but it is possible that youth viewing of advertising aimed at their parents may have influenced family functioning in some way such as decreasing youth resistance to parental monitoring activities. On balance, however, we thought it far more likely that parental monitoring and family functioning would shape youth cognitions about marijuana use. We did not control on parent-child talk because of concerns that some of this talk may have been initiated by the youth after viewing Media Campaign ads and thus be causally posterior to exposure.

Note that many of these parental attributes may be causally prior to *parental* exposure to Media Campaign advertising, but that this is irrelevant for study of the association of youth cognitions with direct *youth* exposure. More complex analyses will be undertaken in the final report to try to determine whether there is a causal relationship between parental exposure and youth outcomes.

### C.3.3 Admissible Pool of Parent Confounders

The committee judged that the following variables properly belong to the pool of admissible potential confounders:

1. Race ethnicity
1. Parent gender
1. Parent age
1. Parental income
1. Parental marital status
1. Parental religiosity
1. Parent has a child aged 9 to 11
1. Parent has a child aged 12 to 13
1. Parent has a child aged 14 to 18
1. Neighborhood characteristics
1. Urbanity
1. Parental use of the internet
1. Parental consumption of newspapers

1. Parental consumption of magazines
1. Parental consumption of TV
1. Parental radio consumption
1. Parental consumption of specific cable channels targeted by the Media Campaign
1. The primary language in which parents watch TV
1. Parental alcohol use
1. Parental tobacco use
1. Parental prior or current use of hard drugs
1. Parental prior or current use of marijuana
1. Parental prior or current use of inhalants
1. Availability of cable or satellite TV in the household

Parental perceptions of family togetherness were excluded since it was felt that it is too close to some of the outcome measures such as parent-child talk. It was felt that, if the Media Campaign is effective in increasing parent-child conversation and activity (as it was meant to), these could actually change parental perceptions of family togetherness.

## C.3.4 Confounders as Moderators

A moderator is a characteristic or predisposition that makes respondents more or less susceptible to the Media Campaign. Moderators may cause the effects of the Media Campaign to be different in different subgroups of the population. In this case, there are interactions of Campaign effects with preexisting factors (the moderators). In this report the moderators that are examined for youth are:

- Age of youth
- Gender of youth
- Race of youth
- Hispanic ethnicity of youth
- Urbanity of home neighborhood
- Natural sensation-seeking tendencies of youth

For parents, the moderators examined in this report are:

- Age of youth
- Gender of youth
- Race of youth
- Hispanic ethnicity of youth
- Urbanity of home neighborhood
- Gender of responding parent
- Education of responding parent

## C.4 Summarization of Confounders

There were too many variables in the pool of admissible potential confounders to remove the effects of each individually. Instead, we summarized the information from the pool that tested as relevant. The summarization method is called propensity scoring. The method was introduced by Rosenbaum and Rubin (1983) and is widely used to analyze observational studies (D'Agostino, 1998). It can handle a large number of confounding variables. It is not necessary to develop complex models for all outcome variables, which is an advantage of this method over some of the alternative adjustment methods available. Exposure is conceptualized as a chance event. The probability distribution of exposure varies across people, (i.e., one person may have a high probability of achieving high exposure while others may have only moderate or low chance of doing the same). However, it is assumed that everyone has some chance of achieving every value of exposure. This rules out the existence of subgroups that are constrained to a sub-range of the possible values of exposure.

The following discussion starts with a general overview of propensity scoring followed by an examination of the propensity scoring's "balance"—the extent to which the counterfactual projections of population means for the confounding variables vary across exposure levels. The remainder of Section C.4 looks first at the impact of the counterfactual projections on effective sample sizes. It then presents the four cross-sectional models that were fitted on the combined data from Waves 1, 2, 3, and 4—one each for the youth general exposure index, the youth recall aided exposure index, the parent general exposure index, and the parent recall aided exposure index, followed by the four stable exposure models and the four lagged effect models.

### C.4.1 Propensity Scoring

Within the group of individuals who have the same exposure propensity, associations between outcome and exposure are free of confounding. This is as if exposure had been randomly assigned to individuals as in a designed experiment. An individual's exposure propensity is estimated as his or her propensity score. Since there are two primary measures of exposure used in this report, two propensity scores were estimated, one for each measure of exposure. An individual's propensity is estimated in terms of confounding variables by complex statistical methods.

Propensity scoring frees the regression modeling process from its usual limitation of reliance on a small number of covariates and simplistic functional forms (e.g., linear main effects only). Rather, a complex model with interactions and higher-order terms can be fit at the propensity scoring stage without concern about overparameterization, since the goal is simply to obtain the best estimated probability of group assignment (in this case to exposure level) from the observed covariates. When subsequently included in the regression model, the propensity score carries all the information from the complex covariate model in a single variable, consuming only one degree of freedom. It also avoids the potentially adverse effects of multicollinearity on the stability of the estimates, regardless of the degree of correlation that exists among the covariates. Finally, propensity score technology can accommodate reasonable numbers of missing observations in the covariates, so fewer cases are lost in analytic procedures requiring complete cases for inclusion.

Despite these advances over traditional regression models, propensity scores have limitations. Like traditional methods for removing group nonequivalence, propensity score methods can adjust only for confounding covariates that are observed and measured. This is always a limitation of nonrandomized studies compared with randomized studies, where the randomization tends to

balance the distribution of all covariates, observed and unobserved. However, tests can be devised to determine the robustness of the conclusions to potential influences of unobserved covariates. Such sensitivity analyses suppose that a relevant but unobserved covariate has been left out of the propensity score model. By explicating how this hypothetical unmeasured covariate is related to treatment assignment and outcome, one can estimate how the treatment effect that adjusts for it might change if such a covariate were available for adjustment. Moreover, propensity scores appear to be more robust to certain types of specification error than standard methods. In a simulation to investigate the relative influence of specification error in propensity scores versus regression models, Drake (1993) found that propensity scores are as vulnerable as standard methods to bias from omitted variables, but less vulnerable to bias from variables that are included but in the wrong functional form (e.g., linear rather than quadratic). A second limitation of propensity score methods—that they require reasonably large samples to support the subclassification—will not be a factor here because reasonably large samples are available. Additional concerns have been raised about the effectiveness of propensity scores for multivariate matching, but they are not being proposed for that purpose here.

Standard propensity score methods assume that there are only two levels of exposure. However, in our set up, exposure is a three- or four-level variable. For this more complex problem, the method suggested by Joffe and Rosenbaum (1990) was used. With this method, an ordinal logit model is fit for each index. The structure of this model is

$$\ln\left(\frac{\sum_{j\le k} p_{ij}}{1-\sum_{j\le k} p_{ij}}\right)=a_k + X_i\beta\,.$$

Here $p_{ij}$ is the propensity of the *i*-th subject for exposure level *j*, $X_i$ denotes the vector of confounder scores for the same subject, $\alpha_k$ is a threshold parameter for the *k*-th exposure level, and $\beta$ is a vector of slope parameters with one component for every confounder retained in the model. The point of the modeling exercise is to identify which of the admissible potential confounders are actually predictive of exposure and then to estimate the vector of slope parameters for those predictors. To fit this model, we used a stepwise variable selection procedure in SAS on the set of potential confounders. (The sampling weights were ignored in fitting the model.)

Once the models had been fit, the next step was to use the model to remove the effects of the confounding variables from the causal analysis. This was done by following a suggestion by Imbens (2000) with some innovations. The basic suggestion of Imbens was to use the estimated propensities to calculate the expected response across the entire sample, which would be expected in the counterfactual event that everyone in the sample had received the same exposure level. This could be achieved with the estimator

$$\hat{y}_{Ck} = \sum_i \frac{\delta_{ik} y_i}{\hat{p}_{ik}}\,,$$

where $\delta_{ik}$ is an indicator variable for the *i*-th case having exposure level *k*, i.e.,

$$\delta_{ik} = \begin{cases} 1 & \text{if the } i\text{-th individual has observed exposure at level } k \\ 0 & \text{else} \end{cases}$$

and $\hat{p}_{ik}$ is the estimated propensity the *i*-th individual has for exposure level *k.* Note that, for each *i*, $\sum_k \hat{p}_{ik} = 1$ for every i.

One innovation for this report was to project the expected response to the entire eligible population by using the sampling weights. This is important in this study given the differential probabilities of selection for youth and parents, depending on family composition. As noted in Appendix A, youth aged 14 to 18 had a higher probability of selection if they had siblings in the 12 to 13 or 9 to 11 brackets, all youth had a lower probability of selection if they had a sibling in the same age bracket, and married parents had lower probabilities of selection than single parents. Also, there is variation in the probability of response to the survey that is reflected in the sampling weights. Using the sampling weights, the counterfactual estimator of response on variable *y* to exposure *k* would be

$$\hat{Y}_{Ck} = \sum_i \frac{\delta_{ik} y_i w_i}{\hat{p}_{ik}},$$

where $w_i$ is the sampling weight for the *i*-th respondent, adjusted for nonresponse and poststratified to population controls. However, it was found that this estimator was unstable and did not balance the covariates very well. Much better results were obtained by smoothing and calibrating the propensities that were estimated by the ordinal logit regression model. The smoothing and calibration was done as follows.

First, the observations were ordered according to the value of $X_i\hat{\beta}$ obtained from the fitted ordinal logit model. The ordered observations were then split into five approximately equal sized groups. Within each group, smoothed and calibrated propensities $\hat{p}$ were calculated according to the formula:

$\tilde{p}_{ik} = \dfrac{\sum_{j \in G_i} \delta_{jk} w_j}{\sum_{j \in G_i} w_j}$, where $G_i \; for \; i \in \{1,2,3,4,5\}$ denotes the group to which observation *i* belongs.

These propensities are smoothed in the sense that there are only five distinct values for each exposure level instead of having a different value for every study subject as is the case with the propensities estimated by the ordinal logit model. These propensities are calibrated in the sense that when they are used to estimate the size of the total population based only on the sample that received a particular exposure level, they yield the same population estimate as is yielded by the total sample. This property is useful in terms of reducing the variance on comparisons of outcomes between exposure levels. The calibration property can be expressed mathematically as

$$\sum_i \frac{\delta_{ik} w_i}{\tilde{p}_{ik}} = \sum_i w_i \;\; \forall \;\; k.$$

Using these smoothed and calibrated propensities and the sampling weights, the counterfactual projection of the average population response on attribute *y* to exposure level *k* is

$$\tilde{Y}_{Ck} = \sum_i \frac{\delta_{ik} y_i w_i}{\tilde{p}_{ik}}.$$

## C.4.2 Assessment of Balance

Because propensity scoring is designed to remove the effects of confounding variables from the association between outcomes and exposures, the counterfactual projections of population means for the confounding variables should not vary across the exposure levels. This property is referred to as

balance. If a confounder has been successfully balanced, then it will have the same counterfactual projection across all exposure levels. Mathematically, this condition of balance is expressed as

$$\sum_i \frac{\delta_{ik} x_{ji} w_i}{\tilde{p}_{ik}} = \sum_i x_{ji} w_i \qquad \forall\, j \text{ and } \forall\, k \ .$$

Figure C-B shows the plots of balance for the 46 significant main variables in modeling youth specific exposure. Means at each level of exposure are denoted by "O" and a 3-standard error range is indicated for each. The overall mean is also indicated on each plot. These plots were generated for the 3-level specific exposure index for youth using the Wave 1 exposure and covariates and the longitudinal weights. Given a large number of covariates to be balanced, there is no expectation that the counterfactual means will be exactly equal. However, if any of the differences are large, there is a risk of bias in the causal analyses due to less than complete control of confounding variables. Balance was also tested on age subgroups but because of the large number of graphs these are not presented. The graphs indicate that balance was achieved overall and in the age subgroups.

## C.4.3 Impact of Counterfactual Projections on Effective Sample Sizes

For the youth general exposure example, the design effects due to the variation in propensities are given in Table C-A. They were calculated using the standard Kish approximation. The true effective samples sizes will be smaller because of larger design effects due to variation in the $W_i$ and due to clustering, but this table gives an impression of how much the counterfactual projection reduces effective sample sizes. The counterfactual projections did not considerably increase variances for the groups with medium or high exposure. The increase in variance for the low-exposure group indicated that confounders were identified that successfully predicted who would have low exposure. The result for correcting for self-selection is a 34 percent reduction in the effective sample size or a 25 percent increase in variances. This was judged to be a good exchange between variance and potential bias.

**Table C-A. Design effects and sample sizes by exposure level**

| Exposure level | Nominal sample size | Design effect | Effective sample size |
|---|---|---|---|
| 1 | 970 | 1.34 | 724 |
| 2 | 1,018 | 1.02 | 1,001 |
| 3 | 2,218 | 1.08 | 2,055 |

**Figure C-B. Plots of balance for lagged youth specicic exposure**

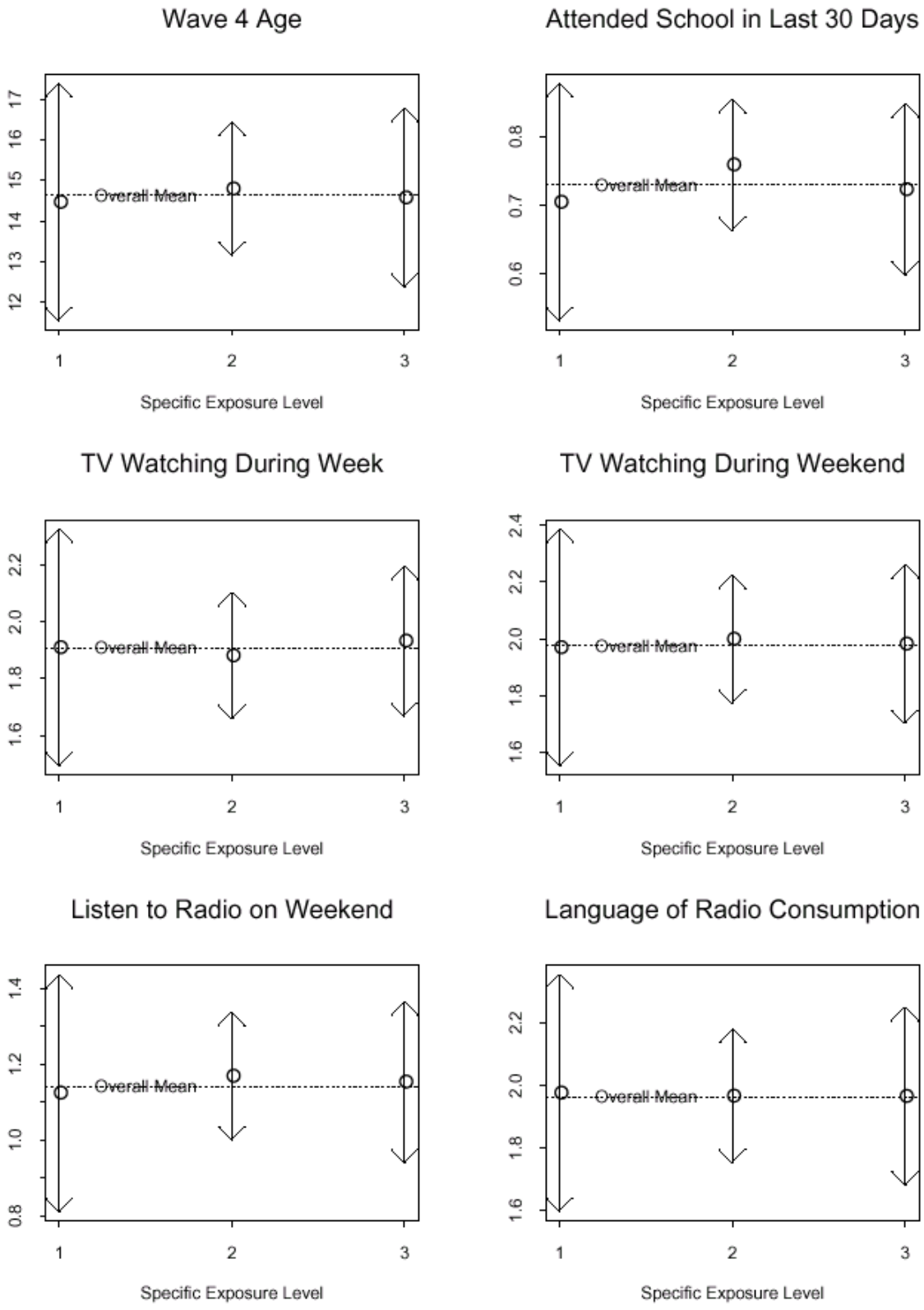


**Figure C-B. Plots of balance for lagged youth specicic exposure (continued)**

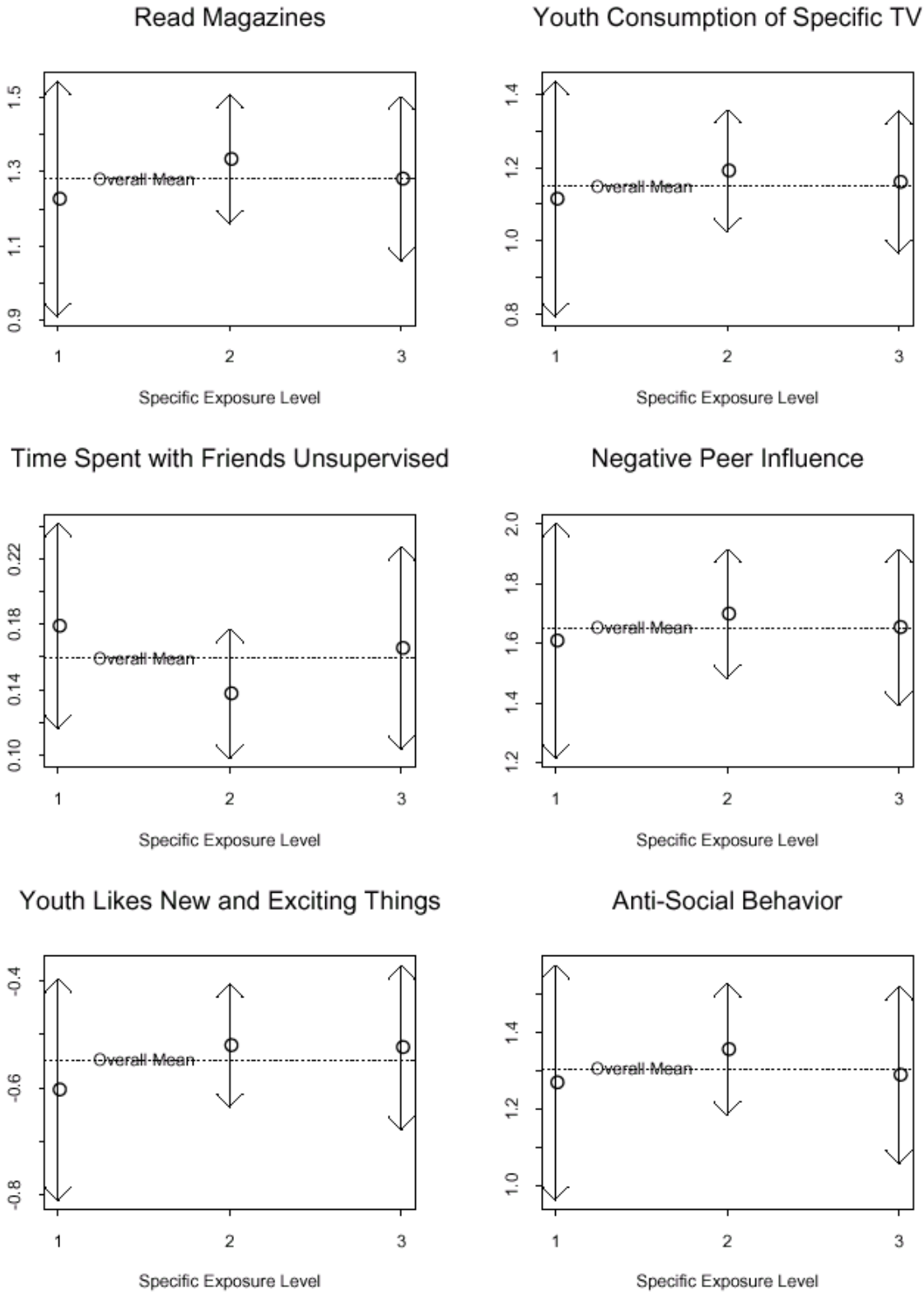


**Figure C-B. Plots of balance for lagged youth specicic exposure (continued)**

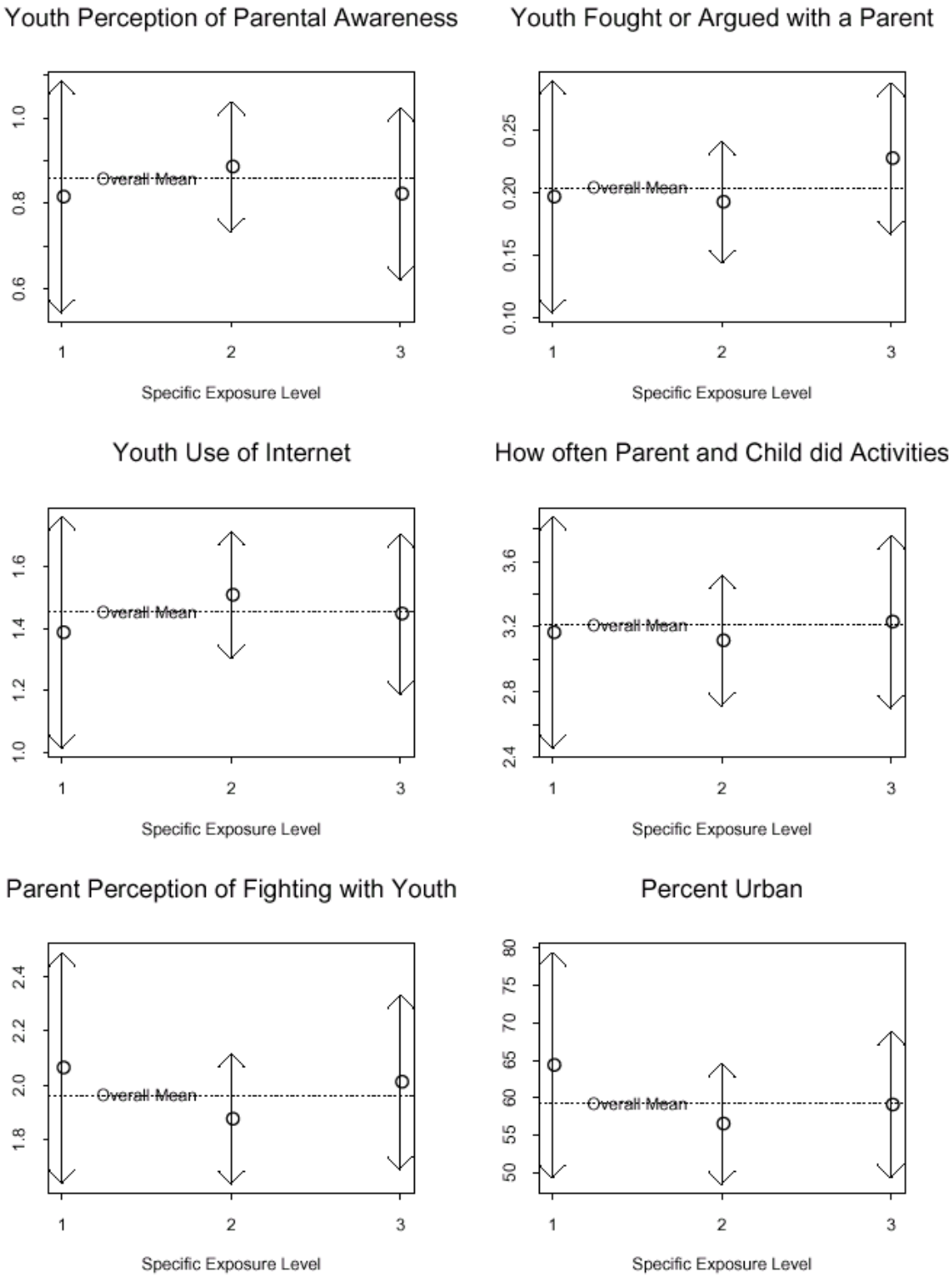


**Figure C-B. Plots of balance for lagged youth specicic exposure (continued)**

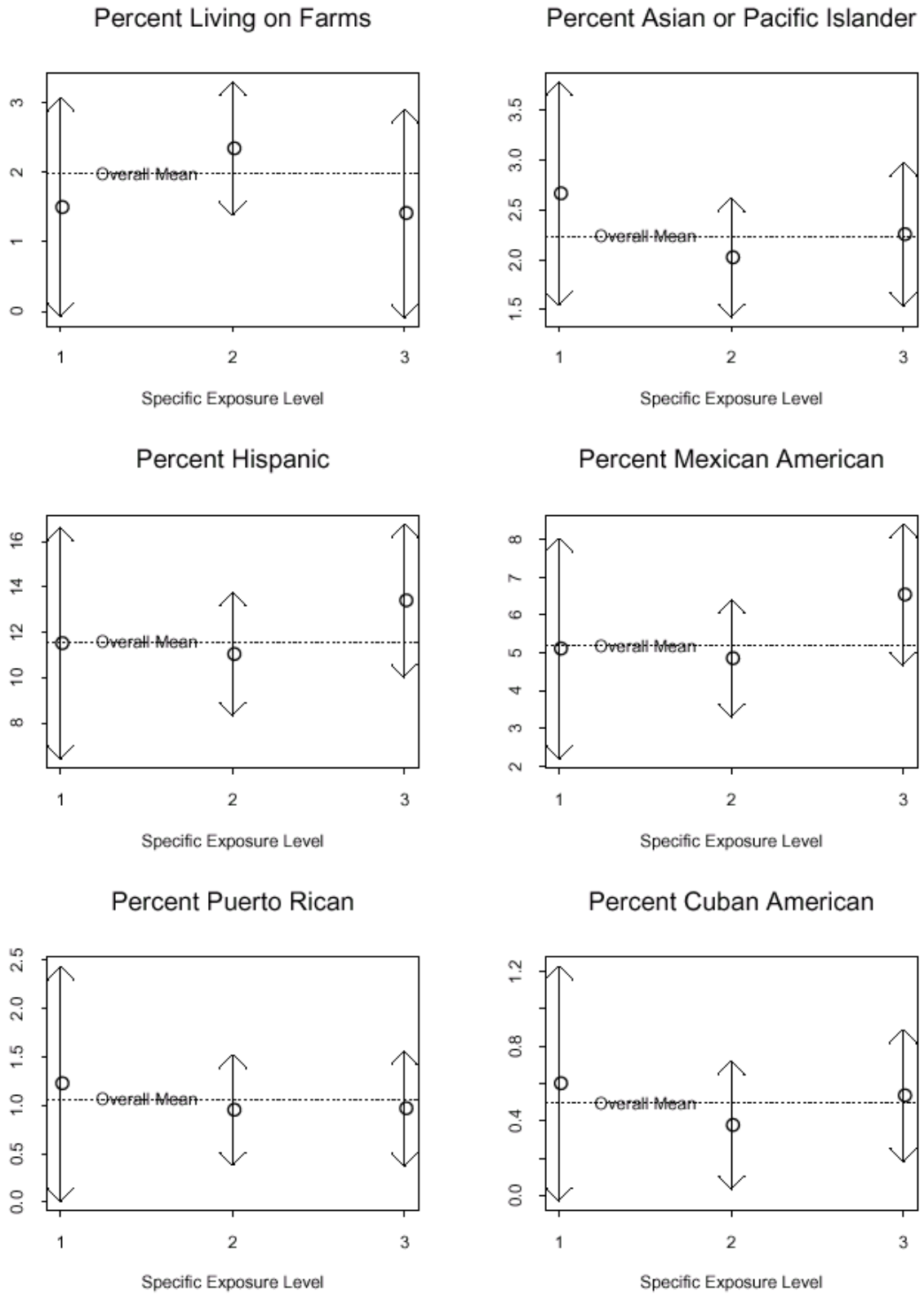


**Figure C-B. Plots of balance for lagged youth specicic exposure (continued)**

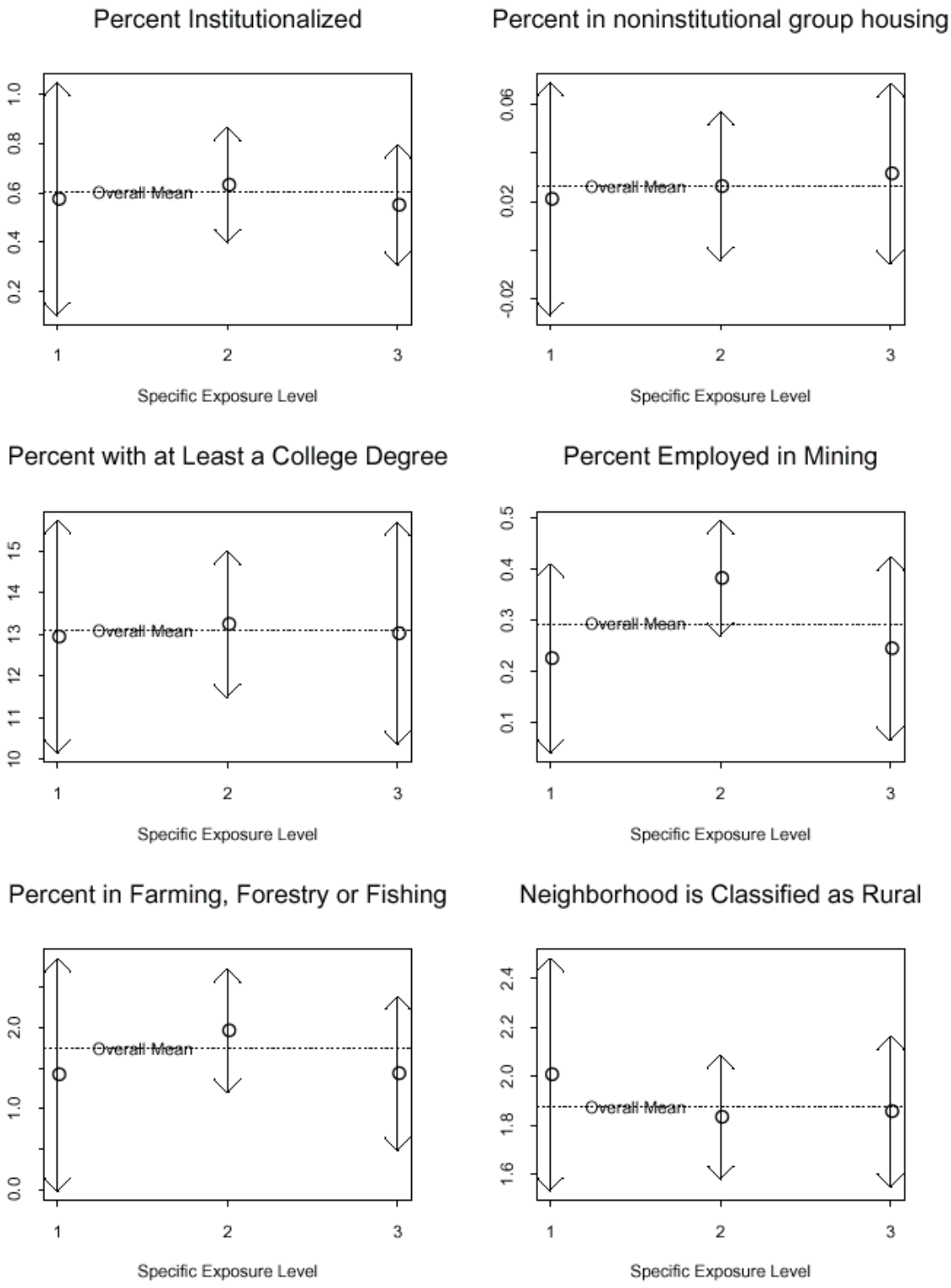


**Figure C-B. Plots of balance for lagged youth specicic exposure (continued)**

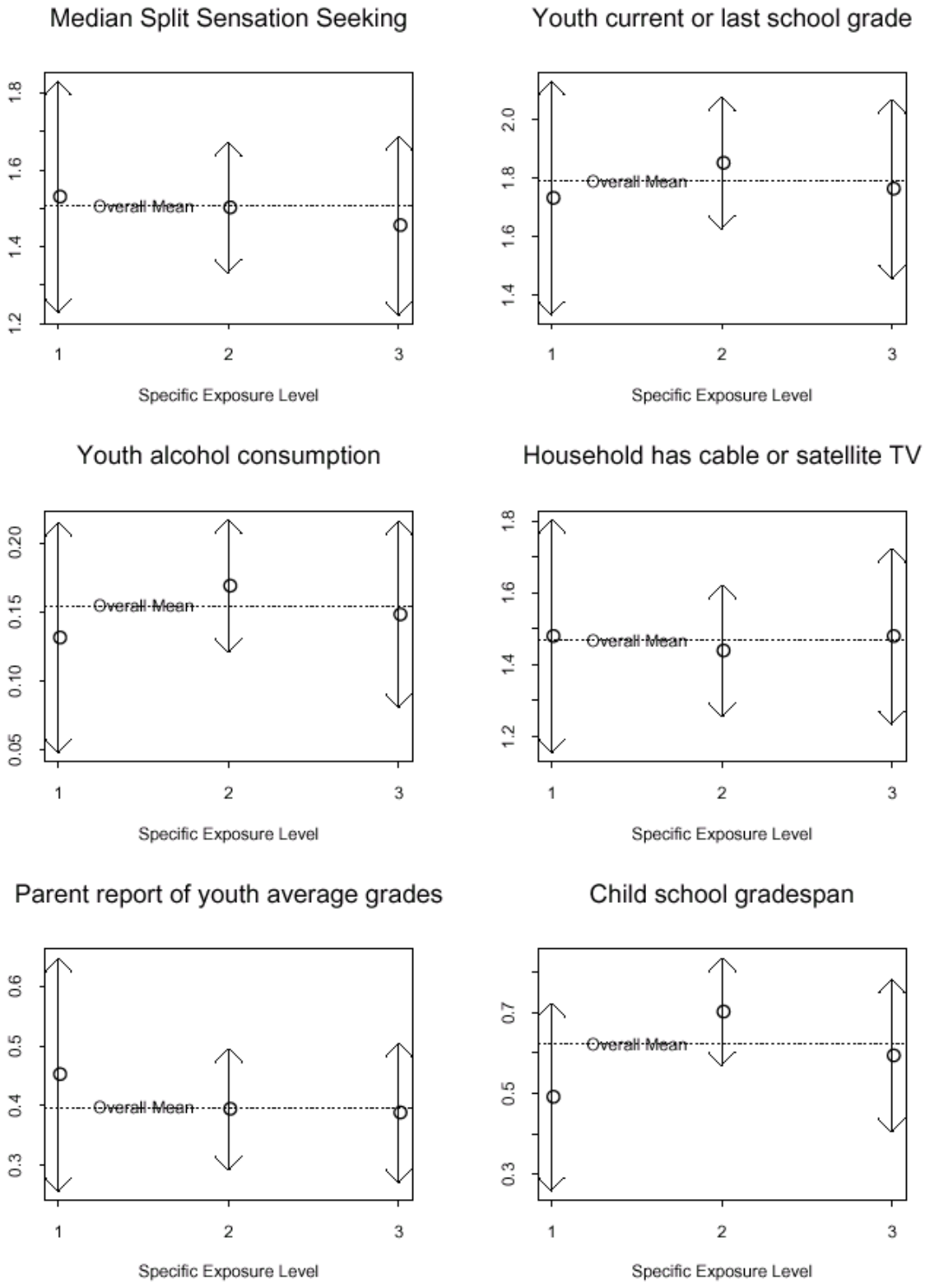


**Figure C-B. Plots of balance for lagged youth specicic exposure (continued)**

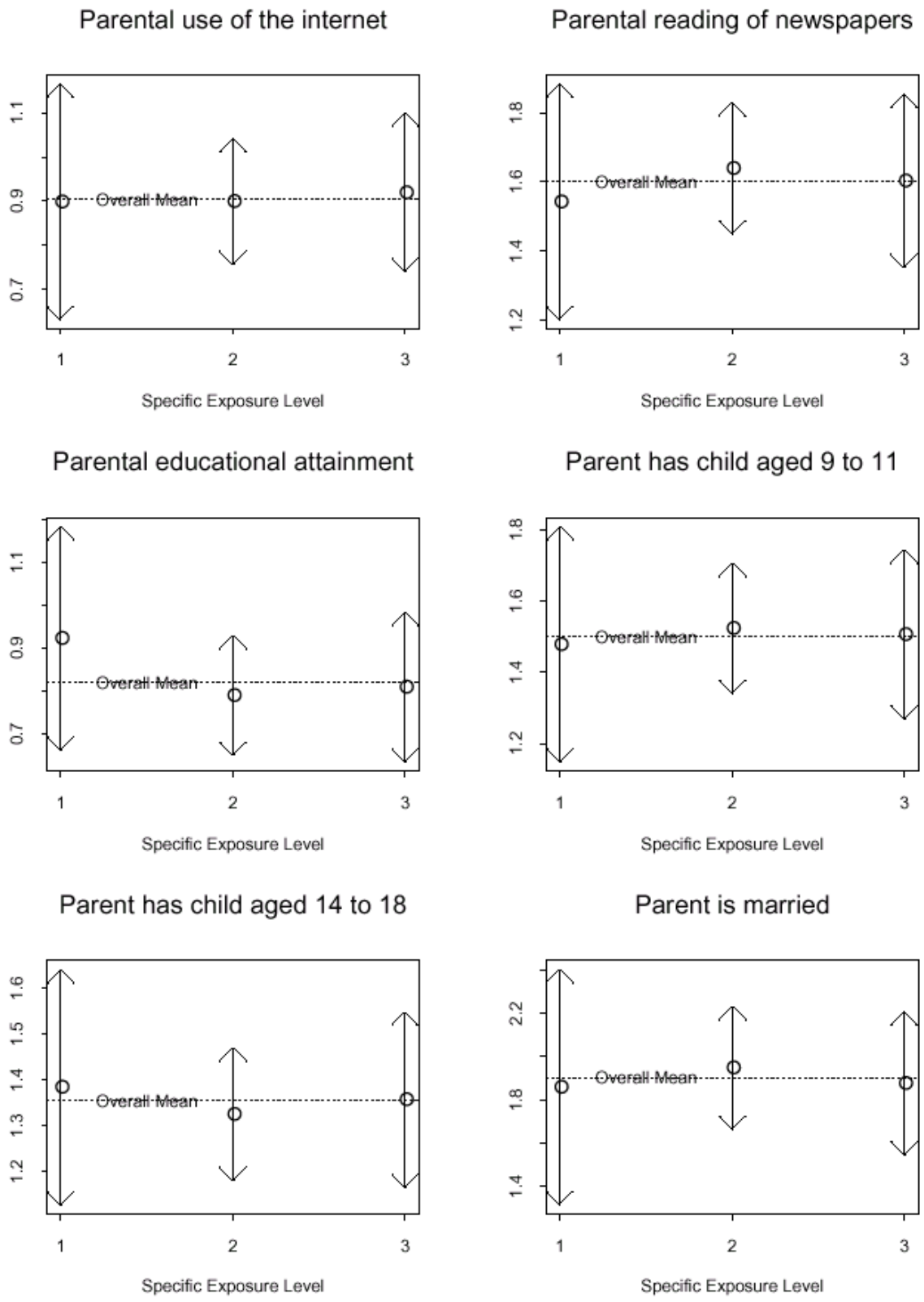
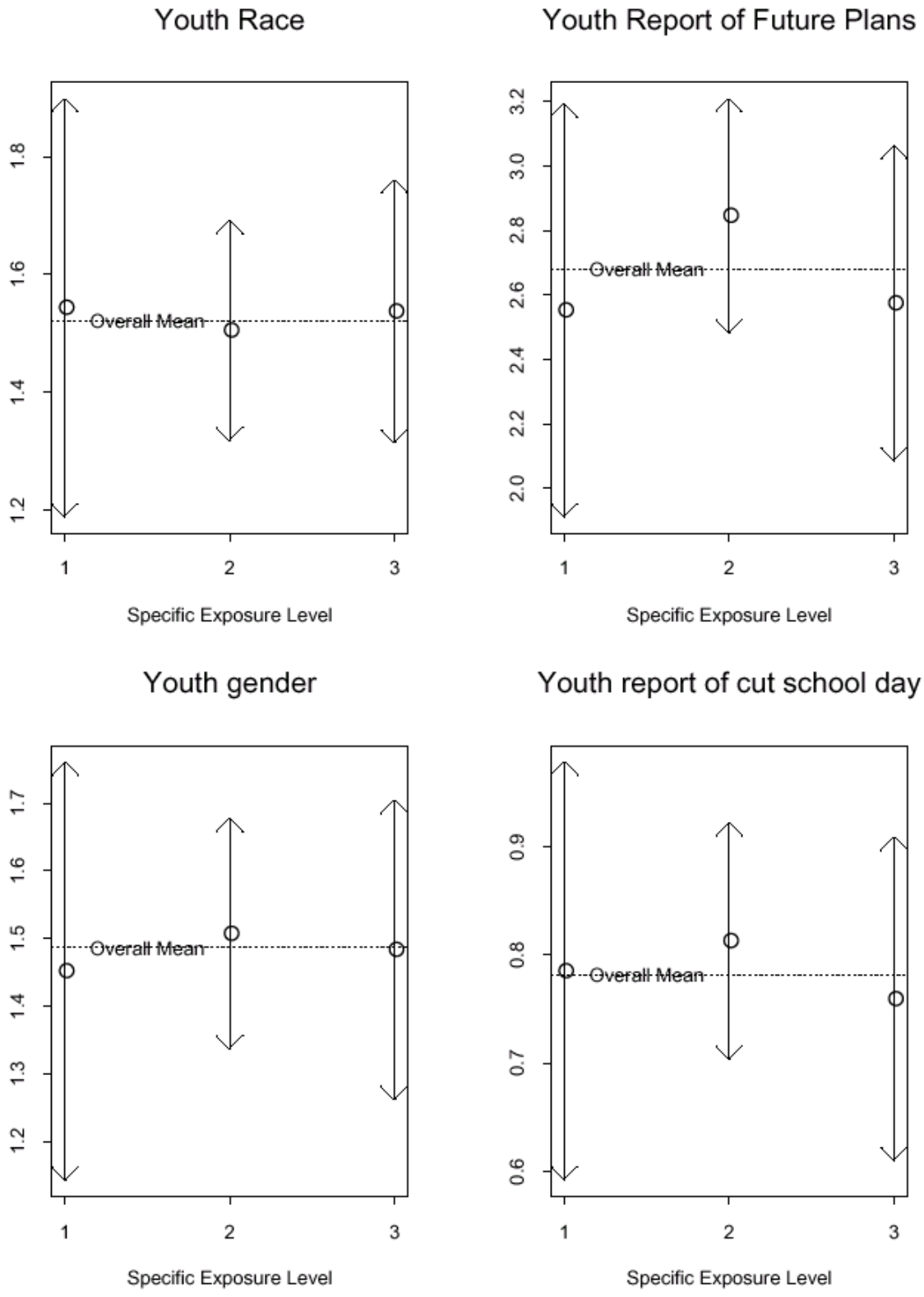


**Figure C-B. Plots of balance for lagged youth specicic exposure (continued)**



## C.4.4 Detailed Models of Exposure

In this section, we present the models that were fitted on the combined data from Waves 1, 2, 3, and 4. Four cross-sectional models were fitted, one for each type of parent exposure index and one for each type of youth exposure index. The variables that were included as potential confounders for each analysis depend on whether the analysis was for parents or for youth. The detailed list of the potential confounders is given in section 3.2 for parents and section 3.1 for youth.

As noted previously, this Wave 4 semiannual report is the first in which longitudinal data were available for analysis. Two longitudinal analyses were proposed to and approved by NIDA: a stable exposure analysis and a lagged analysis. The stable exposure analysis used the average of Wave 1 and Wave 4 exposure as the exposure measure. The lagged analysis used only the Wave 1 exposure data. To meet the requirements of the two longitudinal analyses, new propensity models had to be fit for each. The stable model used the same confounders as the cross-sectional but predicted the new exposure variable. The lagged model for youth was identical to the cross-sectional model, while the lagged model for parents added Wave 1 outcomes to the confounder pool. The lagged model for youth *would* have added Wave 1 outcomes to the confounder pool, except these were not measured on 9- to 11-year-olds in Wave 1. In all, there were eight longitudinal propensity models: youth stable general exposure, youth stable specific exposure, youth lagged general exposure, youth lagged specific exposure, parent stable general exposure, parent stable specific exposure, parent lagged general exposure, and parent lagged specific exposure.

These reduced models were fit using the stepwise ordinal logit procedure in SAS. No weights were used in the model fitting. A level of 0.05 was set for variables to enter the model. Research on Wave 1 indicated that adding interaction terms and other higher order terms tended to result in overfit models that increased the variance of the counterfactual projections without improving balance. However, a term for Wave was included in the cross-sectional models and this term was allowed to interact with other terms in the model. For the cross-sectional models, the Wave term was important for both parent and youth model of recall-aided exposure index but was not important for youth general index.

For categorical variables, coefficients were produced for each value, so that the sum of the coefficients was zero. A positive coefficient means that subjects with that value of the variable have higher exposure propensity, while a negative coefficient means that subjects have lower exposure propensity. For continuous variables, a positive coefficient means that exposure propensity tends to be higher for subjects with higher values of the variable, while a negative coefficient means the opposite.

### C.4.4.1 Cross-Sectional Model for the Youth General Exposure Index

The cross-sectional ordinal logit model found 23 significant variables in modeling youth general exposure compared to 17 for Wave 3. Three of the variables in the Wave 3 report were no longer significant. This is the only model in which wave did not enter as a significant predictor. The variables, together with their coefficients, are presented below in Table C-B.

**Table C-B. Cross-Sectional Model for youth general exposure index among youth aged 12 to 18 who had never tried marijuana**

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of persons who are urban and live inside urbanized areas in the neighborhood in 1990 | Continuous (0-100) | | 0.0016 | 0.0007 |
| | *Percentage of persons in the neighborhood in 1990 who had at least a bachelor's degree | Continuous (0-100) | | -0.0082 | 0.0028 |
| | *Percentage of persons 16+ employed in mining in 1990 | Continuous (0-100) | | -0.0929 | 0.0287 |
| A27, A28, A28a, A28b | Youth consumption of specific cable channels targeted by the Media Campaign | Continuous range from 0 (did not watch) to 3 (watched several channels often) | | 0.3953 | 0.0512 |
| C34a-d | *Youth likes new and exciting friends and experiences | Continuous range from –2 (dislikes) to +2 (strongly likes) | | -0.198 | 0.0696 |
| C35a, C35b | Youth perception of parental awareness of youth activities and plans | Continuous range from –2 (never aware) to +2 (always aware) | | 0.2064 | 0.0249 |
| A19 | Time youth spends watching TV on an average weekday | 0 | Less than 1 hour | -0.2227 | 0.0461 |
| | | 1 | A few hours | 0.0515 | 0.0363 |
| | | 2 | Several hours | 0.1712 | 0.0522 |
| A20 | Time youth spends watching TV on weekends | 0 | Less than 2 hours | -0.1872 | 0.0469 |
| | | 1 | Several hours | 0.0406 | 0.0362 |
| | | 2 | More than 9 hours | 0.1466 | 0.0513 |
| A22 | *Time youth spends listening to Radio on weekends | 0 | Less than 2 hours | -0.1262 | 0.0405 |
| | | 1 | Several hours | 0.0639 | 0.0435 |
| | | 2 | More than 9 hours | 0.0624 | 0.0604 |
| A23 | *Language of Radio programs heard by youth | 0 | More English than Spanish | 0.2328 | 0.0793 |
| | | 1 | At least as much Spanish as English | -0.2328 | 0.0793 |
| C36c | Youth fought or argued with a parent in the last 30 days | 0 | Never or sometimes | 0.0868 | 0.0327 |
| | | 1 | Often or always | -0.0868 | 0.0327 |
| C34c, C34d | Youth score on sensation seeking tendencies (median split) | 0 | Low | 0.1062 | 0.0276 |
| | | 1 | High | -0.1062 | 0.0276 |
| D27 | Youth use of the Internet | 0 | Rarely | -0.2708 | 0.0412 |
| | | 1 | Sometimes | 0.0501 | 0.0351 |
| | | 2 | Every day | 0.2208 | 0.0392 |
| C30 | *Time spent with friends without adult supervision | 0 | Never or some | -0.1226 | 0.0396 |
| | | 1 | Often or always | 0.1226 | 0.0396 |
| A24 | Youth reading of magazines | 0 | Rarely | -0.2073 | 0.0288 |
| | | 1 | Often | 0.2073 | 0.0288 |
| A8, A9 | Youth current or last school grade | 0 | Primary (Grades 1-6) | -0.2332 | 0.0592 |
| | | 1 | Middle (Grades 7-9) | 0.1105 | 0.0429 |
| | | 2 | High (Grades 10-12) | 0.1227 | 0.0630 |
| A9 (parents) | Household has cable or satellite TV service | 0 | No | 0.0993 | 0.0336 |
| | | 1 | Yes | -0.0993 | 0.0336 |

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| A5 (parents) | Parental reading of newspapers | 0 | Never | 0.0005 | 0.0634 |
| | | 1 | Sometimes | -0.126 | 0.0501 |
| | | 2 | Often | 0.1255 | 0.0419 |
| H6 (parents) | Parental educational attainment | 0 | HS or Less | 0.0777 | 0.0285 |
| | | 1 | Beyond HS | -0.0777 | 0.0285 |
| C10 (parents) | *Parent fought or argued with Child in last 30 days | 1 | Never | 0.0872 | 0.0555 |
| | | 2 | Sometimes | -0.1277 | 0.0479 |
| | | 3 | About half the time | -0.0691 | 0.0697 |
| | | 4 | Often | 0.0189 | 0.0778 |
| | | 5 | Always | 0.0908 | 0.1160 |
| | *Neighborhood is classified as a city in a nonurban area (lower population and density) | 0 | No | -0.0841 | 0.0410 |
| | | 1 | Yes | 0.0841 | 0.0410 |
| | Youth was aged between 16 and 18 | 0 | No | 0.1593 | 0.0490 |
| | | 1 | Yes | -0.1593 | 0.0490 |
| A16 | *Youth plans to graduate from 2-year college | 0 | No | -0.1118 | 0.0414 |
| | | 1 | Yes | 0.1118 | 0.0414 |

*Not significant in Wave 3 report.

### C.4.4.2 Cross-Sectional Model for the Youth Recall-Aided Exposure Index

The cross-sectional model for the youth recall-aided exposure index found 33 significant variables compared to 20 variables in Wave 3. Wave was the most significant variable and four wave interaction terms entered the model. These variables are presented, together with their coefficients, in Table C-C.

### C.4.4.3 Cross-Sectional Model for the Parent General Exposure Index

There were 17 significant variables in the model for parental general exposure compared to 14 in the Wave 3 report. Wave as well as one interaction term entered the model. These and their coefficients are tabulated in Table C-D.

### C.4.4.4 Cross-Sectional Model for the Parent Recall-Aided Exposure Index

The model for parental recall-aided exposure found 27 significant variables. In the Wave 3 report there were 23 significant variables for this model. Wave as well as six wave interaction terms entered the model. These variables and their coefficients are presented in Table C-E.

### C.4.4.5 Stable Model for the Youth General Exposure Index

The stable ordinal logit model found 17 significant variables in modeling youth general exposure. The variables, together with their coefficients, are presented below in Table C-F.

### C.4.4.6 Stable Model for the Youth Recall-Aided Exposure Index

The stable model for the youth recall-aided exposure index found 17 significant variables. These variables are presented, together with their coefficients, in Table C-G.

### C.4.4.7 Stable Model for the Parent General Exposure Index

There were 11 significant variables in the stable model for parental general exposure. These and their coefficients are tabulated in Table C-H.

### C.4.4.8 Stable Model for the Parent Recall-Aided Exposure Index

The stable model for parental recall-aided exposure found eight significant variables. These variables and their coefficients are presented in Table C-I.

### C.4.4.9 Lagged Model for the Youth General Exposure Index

The lagged ordinal logit model found 15 significant variables in modeling youth general exposure. The variables, together with their coefficients are presented below in Table C-J.

**Table C-C. Cross-Sectional Model for youth specific exposure index among youth aged 12 to 18 who had never tried marijuana**

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of persons in the neighborhood in 1990 who were Asian or Pacific Islander | Continuous (0-100) | | -0.0091 | 0.0042 |
| | *Percentage of persons in the neighborhood in 1990 who were Hispanic | Continuous (0-100) | | -0.0031 | 0.0014 |
| | *Percentage of persons in the neighborhood in 1990 who were institutionalized | Continuous (0-100) | | 0.0338 | 0.0125 |
| | *Percentage of persons in the neighborhood in 1990 who lived noninstitutional group quarters | Continuous (0-100) | | 0.1098 | 0.0476 |
| | *Percentage of persons in the neighborhood in 1990 who had at least a bachelor's degree | Continuous (0-100) | | -0.0088 | 0.0029 |
| | *Percentage of persons in the neighborhood 16 and older who were employed in 1990 | Continuous (0-100) | | 0.0119 | 0.0027 |
| | Percentage of persons in the neighborhood 16 and older who were in the labor force but unemployed in 1990 | Continuous (0-100) | | 0.0484 | 0.0107 |
| | *Percentage of persons in the neighborhood 16 and older who were employed in mining in 1990 | Continuous (0-100) | | -0.1256 | 0.0276 |
| | *Percentage of persons in the neighborhood 16 and older with farming, forestry, or fishing occupations in 1990 | Continuous (0-100) | | -0.0177 | 0.0071 |
| A27, A28, A28a, A28b | Youth consumption of specific cable channels targeted by the Media Campaign | Continuous range from 0 (did not watch) to 3 (watched several channels often) | | 0.4076 | 0.0486 |
| C31a – C31d | Youth association with antisocial peers | Continuous range from 0 (did not associate) to 6 (associated often) | | 0.0689 | 0.0272 |
| | **Wave of data collection | 1 | Wave 1 | -0.3891 | 0.0911 |
| | | 2 | Wave 2 | -0.0847 | 0.1007 |
| | | 3 | Wave 3 | 0.1490 | 0.1082 |
| | | 4 | Wave 4 | 0.3248 | 0.0956 |
| A19 | Time youth spends watching TV on an average weekday | 0 | Less than 1 hour | -0.2718 | 0.0440 |
| | | 1 | A few hours | 0.0661 | 0.0340 |
| | | 2 | Several hours | 0.2057 | 0.0488 |
| A20 | Time youth spends watching TV on weekends | 0 | Less than 2 hours | -0.2272 | 0.0453 |
| | | 1 | Several hours | 0.0257 | 0.0345 |
| | | 2 | More than 9 hours | 0.2014 | 0.0484 |
| A24 | *Youth reading of magazines | 0 | Rarely | -0.0637 | 0.0277 |
| | | 1 | Often | 0.0637 | 0.0277 |
| C30 | *Time spent with friends without adult supervision | 0 | Never or some | -0.1255 | 0.0382 |
| | | 1 | Often or always | 0.1255 | 0.0382 |
| C34a-d | *Youth likes new and exciting friends and experiences | Continuous range from –2 (dislikes) to +2 (strongly likes) | | 0.0626 | 0.0279 |
| C35a, C35b | *Youth perception of parental awareness of youth activities and plans | Continuous range from –2 (never aware) to +2 (always aware) | | 0.0656 | 0.0239 |
| C36c | Youth fought or argued with a parent in the last 30 days | 0 | Never or sometimes | -0.0752 | 0.0303 |
| | | 1 | Often or always | 0.0752 | 0.0303 |

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| D27 | Youth use of the Internet | 0 | Rarely | -0.1811 | 0.0406 |
| | | 1 | Sometimes | -0.0388 | 0.0331 |
| | | 2 | Every day | 0.2199 | 0.0377 |
| C5c (parents) | How often parents did activities with children in past week | 1 | Not at all | -0.0488 | 0.0563 |
| | | 2 | Once | 0.0642 | 0.0564 |
| | | 3 | Twice | 0.1674 | 0.0513 |
| | | 4 | 3 times | 0.0138 | 0.0583 |
| | | 5 | 4 or 5 times | -0.1216 | 0.0603 |
| | | 6 | 6 or 7 times | 0.0327 | 0.0895 |
| | | 7 | More than 7 times | -0.1077 | 0.0838 |
| A9 (parents) | Household has cable or satellite TV service | 0 | No | 0.1081 | 0.0320 |
| | | 1 | Yes | -0.1081 | 0.0320 |
| H1, H2 (parents) | Youth current or last school grade | 0 | Primary (Grades 1-6) | 0.1032 | 0.0371 |
| | | 1 | Middle (Grades 7-9) | -0.0848 | 0.0438 |
| | | 2 | High (Grades 10-12) | -0.0185 | 0.0404 |
| F9 (parents) | *Parental use of the internet | 0 | Rarely | -0.0416 | 0.0361 |
| | | 1 | Sometimes | -0.0516 | 0.0438 |
| | | 2 | Every day | 0.0932 | 0.0405 |
| | *Parent has child aged 12 to 13 | 0 | No | -0.0703 | 0.0295 |
| | | 1 | Yes | 0.0703 | 0.0295 |
| | Youth gender | 0 | Male | -0.1480 | 0.0252 |
| | | 1 | Female | 0.1480 | 0.0252 |
| H5 (parents) | *Parent is married | 0 | No | 0.0547 | 0.0278 |
| | | 1 | Yes | -0.0547 | 0.0278 |
| A3, A4, A5 | *Youth was Hispanic | 0 | No | -0.2159 | 0.0458 |
| | | 1 | Yes | 0.2159 | 0.0458 |
| | Youth was aged between 16 and 18 | 0 | No | 0.1324 | 0.0352 |
| | | 1 | Yes | -0.1324 | 0.0352 |
| A20 | **Interaction of Wave and Time youth spends watching TV on weekend1s | 1 1 | Wave 1 – A Little | 0.0788 | 0.0640 |
| | | 1 2 | Wave 1 – Some | -0.0950 | 0.0530 |
| | | 1 3 | Wave 1 – A Lot | 0.0161 | 0.0670 |
| | | 2 1 | Wave 2 – A Little | -0.0441 | 0.0731 |
| | | 2 2 | Wave 2 – Some | 0.1053 | 0.0611 |
| | | 2 3 | Wave 2 – A Lot | -0.0612 | 0.0778 |
| | | 3 1 | Wave 3 – A Little | 0.1506 | 0.0710 |
| | | 3 2 | Wave 3 – Some | -0.0846 | 0.0605 |
| | | 3 3 | Wave 3 – A Lot | -0.0660 | 0.0774 |
| | | 4 1 | Wave 4 – A Little | -0.1854 | 0.0655 |
| | | 4 2 | Wave 4 – Some | 0.0743 | 0.0555 |
| | | 4 3 | Wave 4 – A Lot | 0.1111 | 0.0727 |
| | **Interaction of Wave and Urbanicity of neighborhood (continuous from 0 to 4) | 1 | Wave 1 | -0.0535 | 0.0299 |
| | | 2 | Wave 2 | -0.0456 | 0.0335 |
| | | 3 | Wave 3 | 0.0850 | 0.0340 |
| | | 4 | Wave 4 | 0.0141 | 0.0318 |

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| F9 (parents) | **Interaction of Wave and Parental use of the internet | 1 0 | Wave 1 Rarely use | -0.0174 | 0.0551 |
| | | 1 1 | Wave 1 Sometimes use | -0.0601 | 0.0706 |
| | | 1 2 | Wave 1 Often use | 0.0775 | 0.0589 |
| | | 2 0 | Wave 2 Rarely use | -0.1079 | 0.0625 |
| | | 2 1 | Wave 2 Sometimes use | 0.0140 | 0.0804 |
| | | 2 2 | Wave 2 Often use | 0.0939 | 0.0686 |
| | | 3 0 | Wave 3 Rarely use | 0.1052 | 0.0616 |
| | | 3 1 | Wave 3 Sometimes use | 0.1481 | 0.0798 |
| | | 3 2 | Wave 3 Often use | -0.2533 | 0.0691 |
| | | 4 0 | Wave 4 Rarely use | 0.0201 | 0.0561 |
| | | 4 1 | Wave 4 Sometimes use | -0.1021 | 0.0717 |
| | | 4 2 | Wave 4 Often use | 0.0820 | 0.0639 |
| A3, A4, A5 | **Interaction of Wave and whether youth was Hispanic | 1 0 | Wave 1 Non-Hispanic | 0.1014 | 0.0588 |
| | | 1 1 | Wave 1 Hispanic | -0.1014 | 0.0588 |
| | | 2 0 | Wave 2 Non-Hispanic | -0.0229 | 0.0664 |
| | | 2 1 | Wave 2 Hispanic | 0.0229 | 0.0664 |
| | | 3 0 | Wave 3 Non-Hispanic | -0.2186 | 0.0693 |
| | | 3 1 | Wave 3 Hispanic | 0.2186 | 0.0693 |
| | | 4 0 | Wave 4 Non-Hispanic | 0.1401 | 0.0617 |
| | | 4 1 | Wave 4 Hispanic | -0.1401 | 0.0617 |

*Not significant in Wave 3 report. **Could not enter the Wave 3 model.

**Table C-D. Cross-Sectional Model for parent general exposure index among all parents of youth aged 9 to 18**

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | *Percentage of persons who are urban and live inside urbanized areas in the neighborhood in 1990 | Continuous (0-100) | | 0.0017 | 0.0005 |
| | Percentage of Cuban Americans in the neighborhood in 1990 | Continuous (0-100) | | -0.0148 | 0.0054 |
| | *Percentage of persons in the neighborhood in 1990 who had at least a bachelor's degree | Continuous (0-100) | | -0.0093 | 0.0035 |
| | * Percentage of persons in the neighborhood 16 and older who were in the labor force but unemployed in 1990 | Continuous (0-100) | | 0.0328 | 0.0089 |
| | * Percentage of households with income above $75000 per year in the neighborhood in 1990 | Continuous (0-100) | | -0.0059 | 0.0030 |
| A9a, A9b | Parental viewing of BET and Spanish-language cable channels in the past 30 days | 0 | None | -0.3753 | 0.0368 |
| | | 1 | Less than 15 days | -0.0256 | 0.0397 |
| | | 2 | 15 or more days | 0.4009 | 0.0543 |
| F9 | *Parental use of the internet | 0 | Rarely | 0.0793 | 0.0325 |
| | | 1 | Sometimes | -0.0262 | 0.0403 |
| | | 2 | Every day | -0.0531 | 0.0358 |
| A3, A4 | Time parents spend listening to radio per week | 0 | Less than 1 hour | -0.2081 | 0.0355 |
| | | 1 | A few hours | 0.0117 | 0.0301 |
| | | 2 | Several hours | 0.1964 | 0.0338 |
| A6 | Parental reading of magazines | 0 | Never | -0.2706 | 0.0553 |
| | | 1 | Sometimes | -0.0484 | 0.0360 |
| | | 2 | Often | 0.3191 | 0.0377 |
| H5 | *Parent is married | 0 | No | -0.0606 | 0.0262 |
| | | 1 | Yes | 0.0606 | 0.0262 |
| A5 | Parental reading of newspapers | 0 | Never | -0.1736 | 0.0571 |
| | | 1 | Sometimes | -0.0735 | 0.0442 |
| | | 2 | Often | 0.2471 | 0.0379 |
| H7, H8 | *Influence of religion on parents | 0 | Low | 0.0644 | 0.0294 |
| | | 1 | High | -0.0644 | 0.0294 |
| G1, G2 | Parental smoking behavior | 0 | Never smoked | -0.0318 | 0.0337 |
| | | 1 | Not recently or Rarely | -0.0786 | 0.0307 |
| | | 2 | Twice or more per day | 0.1105 | 0.0355 |
| A1, A2 | Time parents spend watching TV per week | 0 | Less than 1 hour | -0.4562 | 0.0618 |
| | | 1 | A few hours | 0.0873 | 0.0381 |
| | | 2 | Several hours | 0.3689 | 0.0398 |
| H5 | Parent is a widow | 0 | No | -0.1816 | 0.0813 |
| | | 1 | Yes | 0.1816 | 0.0813 |
| | Wave of Data Collection | 1 | Wave 1 | 0.1086 | 0.0397 |
| | | 2 | Wave 2 | -0.0249 | 0.0442 |
| | | 3 | Wave 3 | 0.0237 | 0.0437 |
| | | 4 | Wave 4 | -0.1074 | 0.0435 |

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Interaction of Wave and whether Parent is Married | 0 1 | Wave 1 Not Married | -0.0652 | |
| | | 0 2 | Wave 2 Not Married | 0.0455 | |
| | | 0 3 | Wave 3 Not Married | 0.1019 | |
| | | 0 4 | Wave 4 Not Married | | |
| | | 1 1 | Wave 1 Not Married | | |
| | | 1 2 | Wave 2 Not Married | | |
| | | 1 3 | Wave 3 Not Married | | |
| | | 1 4 | Wave 4 Not Married | | |

*Not significant in Wave 3 report.

**Could not enter the Wave 3 model.

**Table C-E. Cross-Sectional Model for parent specific exposure index among all parents of youth aged 9 to 18**

| Quex Items | Description of variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of American Indian, Eskimo and Aleut persons in the neighborhood in 1990 | | Continuous(0-100) | 0.0101 | 0.00386 |
| | *Percentage of other Hispanic persons in the neighborhood in 1990 | | Continuous(0-100) | 0.00509 | 0.0191 |
| | *Percentage of persons in group quarters in the neighborhood in 1990 | | Continuous(0-100) | 0.0885 | 0.0378 |
| | *Percentage of foreign-born non-citizens in the neighborhood in 1990 | | Continuous(0-100) | -0.0149 | 0.00315 |
| | *Percentage high-school dropouts in the neighborhood in 1990 | | Continuous(0-100) | 0.0117 | 0.00289 |
| | *Percentage of persons 16+ working in manufacturing in the neighborhood in 1990 | | Continuous(0-100) | 0.0124 | 0.00352 |
| | Percentage of persons 16+ working in farming & fishing in the neighborhood in 1990 | | Continuous(0-100) | 0.0236 | 0.00671 |
| | Percentage of households with income above $75000 per year in the neighborhood in 1990 | | Continuous(0-100) | -0.0084 | 0.00234 |
| | *Percentage of large structured Hus in the neighborhood in 1990 | | Continuous(0-100) | -0.00795 | 0.00244 |
| | Neighborhood is classified as a city in a nonurban area (lower population and density) | 0 | No | -0.0994 | 0.0284 |
| | | 1 | Yes | 0.0994 | 0.0284 |
| A9a, A9b | Parent viewing of BET and Spanish-language cable channels in the past 30 days | 0 | None | -0.2414 | 0.0372 |
| | | 1 | Less than 15 days | 0.0359 | 0.0383 |
| | | 2 | 15 or more days | 0.2055 | 0.055 |
| A3, A4 | Time Parents spend listening to radio per week | 0 | Less than 1 hour | -0.2178 | 0.0345 |
| | | 1 | A few hours | 0.0146 | 0.029 |
| | | 2 | Several hours | 0.2032 | 0.0322 |
| H7, H8 | *Influence of religion on parents | 0 | Low | 0.0774 | 0.0293 |
| | | 1 | High | -0.0774 | 0.0293 |
| G1, G2 | Parental smoking behavior | 0 | Never smoked | -0.0619 | 0.0351 |
| | | 1 | Not recently or Rarely | -0.0739 | 0.0297 |
| | | 2 | Twice or more per day | 0.1358 | 0.035 |
| A2b | Language of parental TV viewing | 0 | Spanish | 0.1828 | 0.0555 |
| | | 1 | English | -0.1828 | 0.0555 |
| A1, A2 | Time parents spend watching TV per week | 0 | Less than 1 hour | -0.534 | 0.0616 |
| | | 1 | A few hours | 0.1414 | 0.0377 |
| | | 2 | Several hours | 0.3926 | 0.0393 |

**Table C-E. Cross-Sectional Model for parent specific exposure index among all parents of youth aged 9 to 18**

| Quex Items | Description of variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Parent has child aged 14-18 | 1 | Yes | 0.0469 | 0.0218 |
| | | 2 | No | -0.0469 | 0.0218 |
| | *Gender of parents | 1 | Male | -0.0598 | 0.0233 |
| | | 2 | Female | 0.0598 | 0.0233 |
| H13 | Parental income is missing | 0 | No | 0.1275 | 0.06 |
| | | 1 | Yes | -0.1275 | 0.06 |
| G6 | Parent has ever used marijuana | 0 | No | 0.0439 | 0.0245 |
| | | 1 | Yes | -0.0439 | 0.0245 |
| | **WAVE of Data Collection | 1 | | 0.3979 | 0.1072 |
| | | 2 | | -0.0645 | 0.124 |
| | | 3 | | -0.2439 | 0.118 |
| | | 4 | | -0.0895 | 0.1166 |
| | **Percentage of persons 16+ working in | 1 | Percentage of persons 16+ working in manufacture in wave 1 | -0.0013 | 0.00552 |
| | Manufacturing by wave | 2 | Percentage of persons 16+ working in manufacture in wave 2 | -0.0154 | 0.00593 |
| | | 3 | Percentage of persons 16+ working in manufacture in wave 3 | 0.019 | 0.00621 |
| | | 4 | Percentage of persons 16+ working in manufacture in wave 4 | -0.0023 | 0.0061 |
| | **Influence of religion on parents by wave | 0 1 | Influence of religion on parents is low in wave 1 | -0.1016 | 0.0455 |
| | | 0 2 | Influence of religion on parents is low in wave 2 | -0.0412 | 0.0508 |
| | | 0 3 | Influence of religion on parents is low in wave 3 | -0.0112 | 0.0515 |
| | | 0 4 | Influence of religion on parents is low in wave 4 | 0.154 | 0.0519 |
| | | 1 1 | Influence of religion on parents is high in wave 1 | 0.1016 | 0.0455 |
| | | 1 2 | Influence of religion on parents is high in wave 2 | 0.0412 | 0.0508 |
| | | 1 3 | Influence of religion on parents is high in wave 3 | 0.0112 | 0.0515 |
| | | 1 4 | Influence of religion on parents is high in wave 4 | -0.154 | 0.0519 |
| | **Language of parental TV viewing by wave | 0 1 | Parental TV viewing is in Spanish in wave 1 | 0.355 | 0.0716 |
| | | 0 2 | Parental TV viewing is in Spanish in wave 2 | 0.1012 | 0.088 |
| | | 0 3 | Parental TV viewing is in Spanish in wave 3 | -0.0087 | 0.0778 |
| | | 0 4 | Parental TV viewing is in Spanish in wave 4 | -0.4475 | 0.0763 |

| Quex Items | Description of variable | Values | | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|---|
| | | 1 | 1 | Parental TV viewing is in English in wave 1 | -0.355 | 0.0716 |
| | | 1 | 2 | Parental TV viewing is in English in wave 2 | -0.1012 | 0.088 |
| | | 1 | 3 | Parental TV viewing is in English in wave 3 | 0.0087 | 0.0778 |
| | | 1 | 4 | Parental TV viewing is in English in wave 4 | 0.4475 | 0.0763 |
| | **Time parents spend watching TV per week by wave | 0 | 1 | Parents spend less than 1 hour watching TV in wave 1 | 0.2915 | 0.0972 |
| | | 0 | 2 | Parents spend less than 1 hour watching TV in wave 2 | -0.1608 | 0.1145 |
| | | 0 | 3 | Parents spend less than 1 hour watching TV in wave 3 | 0.2415 | 0.1061 |
| | | 0 | 4 | Parents spend less than 1 hour watching TV in wave 4 | -0.3721 | 0.1034 |
| | | 1 | 1 | Parents spend a few hours watching TV in wave 1 | -0.085 | 0.0603 |
| | | 1 | 2 | Parents spend a few hours watching TV in wave 2 | 0.1837 | 0.0697 |
| | | 1 | 3 | Parents spend a few hours watching TV in wave 3 | -0.135 | 0.0657 |
| | | 1 | 4 | Parents spend a few hours watching TV in wave 4 | 0.0364 | 0.0644 |
| | | 2 | 1 | Parents spend several hours watching TV in wave 1 | -0.2064 | 0.0604 |
| | | 2 | 2 | Parents spend several hours watching TV in wave 2 | -0.0228 | 0.0701 |
| | | 2 | 3 | Parents spend several hours watching TV in wave 3 | -0.1065 | 0.0669 |
| | | 2 | 4 | Parents spend several hours watching TV in wave 4 | 0.3357 | 0.0656 |
| | **Gender of parents by wave | 1 | 1 | Male parent in wave 1 | 0.0818 | 0.0371 |
| | | 1 | 2 | Male parent in wave 2 | 0.0485 | 0.0409 |
| | | 1 | 3 | Male parent in wave 3 | 0.0051 | 0.0411 |
| | | 1 | 4 | Male parent in wave 4 | -0.1354 | 0.0407 |
| | | 2 | 1 | Female parent in wave 1 | -0.0818 | 0.0371 |
| | | 2 | 2 | Female parent in wave 2 | -0.0485 | 0.0409 |
| | | 2 | 3 | Female parent in wave 3 | -0.0051 | 0.0411 |
| | | 2 | 4 | Female parent in wave 4 | 0.1354 | 0.0407 |
| | **Parent has ever used marijuana by wave | 0 | 1 | Parent has never used marijuana in wave 1 | 0.0851 | 0.0359 |
| | | 0 | 2 | Parent has never used marijuana in wave 2 | 0.0161 | 0.0402 |
| | | 0 | 3 | Parent has never used marijuana in wave 3 | 0.0179 | 0.0401 |

| Quex Items | Description of variable | Values | | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|---|
| | | 0 | 4 | Parent has never used marijuana in wave 4 | -0.1191 | 0.0397 |
| | | 1 | 1 | Parent has ever used marijuana in wave 1 | -0.0851 | 0.0359 |
| | | 1 | 2 | Parent has ever used marijuana in wave 2 | -0.0161 | 0.0402 |
| | | 1 | 3 | Parent has ever used marijuana in wave 3 | -0.0179 | 0.0401 |
| | | 1 | 4 | Parent has ever used marijuana in wave 4 | 0.1191 | 0.0397 |

*Not significant in Wave 3 report. **Could not enter the Wave 3 model.

**Table C-F. Stable Model for youth general exposure index among youth aged 12 to 18 at Wave 4 who had never tried marijuana at Wave 1**

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of persons in the neighborhood in 1990 who are urban and live inside urbanized areas | Continuous (0-100) | | 0.00433 | 0.00113 |
| | Percentage of persons in the neighborhood in 1990 who had at least a bachelor's degree | Continuous (0-100) | | -0.0126 | 0.00507 |
| C34a-d | Youth likes new and exciting friends and experiences | Continuous range from –2 (dislikes) to +2 (strongly likes) | | -0.3891 | 0.1451 |
| C35a, C35b | Youth perception of parental awareness of youth activities and plans | Continuous range from –2 (never aware) to +2 (always aware) | | 0.2133 | 0.0472 |
| A19 | Time youth spends watching TV on an average weekday | -1 | Less than 1 hour | -0.2416 | 0.0913 |
| | | 0 | A few hours | 0.1232 | 0.0712 |
| | | 1 | Several hours | 0.1183 | 0.1047 |
| A20 | Time youth spends watching TV on weekends | 0 | Less than 2 hours | -0.2616 | 0.0925 |
| | | 1 | Several hours | 0.0524 | 0.0701 |
| | | 2 | More than 9 hours | 0.2092 | 0.0994 |
| A24 | Youth reading of magazines | 0 | Rarely | -0.163 | 0.0544 |
| | | 1 | Often | 0.163 | 0.0544 |
| C30 | Time spent with friends without adult supervision | 0 | Never or some | 0.1567 | 0.0706 |
| | | 1 | Often or always | -0.1567 | 0.0706 |
| D27 | Youth use of the Internet | 0 | Rarely | -0.2312 | 0.0745 |
| | | 1 | Sometimes | -0.0870 | 0.0680 |
| | | 2 | Every day | 0.3181 | 0.0817 |
| C3 (parents) | How often child hangs out with his/her friends without adult supervision | 0 | Never | 0.1641 | 0.1050 |
| | | 1 | Seldom | 0.2205 | 0.0989 |
| | | 2 | About Half the time | 0.3165 | 0.1499 |
| | | 3 | Often | -0.1756 | 0.1708 |
| | | 4 | Almost Always | | |
| C10 (parents) | Parent fought or argued with Child in last 30 days | 1 | Never | 0.1152 | 0.1122 |
| | | 2 | Sometimes | -0.2494 | 0.0968 |
| | | 3 | About half the time | 0.1866 | 0.1485 |
| | | 4 | Often | -0.1395 | 0.1538 |
| | | 5 | Always | | |
| | Parent has child aged 14 to 18 | 0 | No | -0.1376 | 0.0518 |
| | | 1 | Yes | 0.1376 | 0.0518 |
| H5 (parents) | Parent is living as married | 0 | No | -0.4023 | 0.1778 |
| | | 1 | Yes | 0.4023 | 0.1778 |
| A3, A4, A5 | Youth was Asian or Pacific Islander | 0 | No | -0.1376 | 0.0518 |
| | | 1 | Yes | 0.1376 | 0.0518 |
| A16 | Youth plans to graduate from 2-year college | 0 | No | 0.1786 | 0.0757 |
| | | 1 | Yes | -0.1786 | 0.0757 |

**Table C-G. Stable Model for youth specific exposure index among youth aged 12 to 18 at Wave 4 who had never tried marijuana at Wave 1**

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of persons in the neighborhood in 1990 who are urban and live inside urbanized areas | Continuous (0-100) | | 0.0027 | 0.0010 |
| | Percentage of persons in the neighborhood in 1990 who were Black | Continuous (0-100) | | 0.0096 | 0.0027 |
| | Percentage of persons in the neighborhood in 1990 who were Hispanic | Continuous (0-100) | | -0.0078 | 0.0028 |
| | Persons per room in the neighborhood in 1990 | Continuous (0-100) | | 1.9877 | 0.5038 |
| | Persons per vehicle in the neighborhood in 1990 | Continuous (0-100) | | -0.0720 | 0.0300 |
| A27, A28, A28a, A28b | Youth consumption of specific cable channels targeted by the Media Campaign | Continuous range from 0 (did not watch) to 3 (watched several channels often) | | 0.3578 | 0.0907 |
| A19 | Time youth spends watching TV on an average weekday | -1 | Less than 1 hour | -0.5224 | 0.0753 |
| | | 0 | A few hours | 0.0769 | 0.0632 |
| | | 1 | Several hours | 0.4454 | 0.0868 |
| A6 (parents) | Parental reading of magazines | -1 | Never | 0.1293 | 0.1115 |
| | | 0 | Sometimes | 0.0415 | 0.0751 |
| | | 1 | Often | -0.1708 | 0.0730 |
| C30 | Time spent with friends without adult supervision | 0 | Never or some | 0.2414 | 0.0638 |
| | | 1 | Often or always | -0.2414 | 0.0638 |
| C36c | Youth fought or argued with a parent in the last 30 days | 0 | Never or sometimes | -0.2321 | 0.0585 |
| | | 1 | Often or always | 0.2321 | 0.0585 |
| D27 | Youth use of the Internet | -1 | Rarely | -0.2205 | 0.0701 |
| | | 0 | Sometimes | -0.0601 | 0.0623 |
| | | 1 | Every day | 0.2805 | 0.0735 |
| A9 (parents) | Household has cable or satellite TV service | 0 | No | 0.1290 | 0.0590 |
| | | 1 | Yes | -0.1290 | 0.0590 |
| H1, H2 (parents) | Youth current or last school grade | -1 | Primary (Grades 1-6) | 0.3854 | 0.0860 |
| | | 0 | Middle (Grades 7-9) | 0.0774 | 0.0690 |
| | | 1 | High (Grades 10-12) | -0.4629 | 0.1065 |
| | Parent has child aged 9 to 11 | 0 | No | -0.2529 | 0.0564 |
| | | 1 | Yes | 0.2529 | 0.0564 |
| | Parent has child aged 14 to 18 | 0 | No | -0.1365 | 0.0552 |
| | | 1 | Yes | 0.1365 | 0.0552 |
| | Youth gender | 0 | Male | -0.1539 | 0.0471 |
| | | 1 | Female | 0.1539 | 0.0471 |
| H5 (parents) | Parent is married | 0 | No | 0.1556 | 0.0523 |
| | | 1 | Yes | -0.1556 | 0.0523 |

**Table C-H Stable Model for parent general exposure index among all parents of youth aged 12 to 18 at Wave 4**

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of persons who are urban and live inside urbanized areas in the neighborhood in 1990 | Continuous(0-100) | | -0.00277 | 0.00118 |
| | Percentage of household with income above $75,000 per year | Continuous(0-100) | | 0.0242 | 0.00486 |
| | Talking belief/Attitude index | Continuous | | -0.00242 | 0.000584 |
| A9a, A9b | Parent viewing of BET and Spanish-language Cable channels in the past 30 days | 0 | None | 0.3955 | 0.1078 |
| | | 1 | Less than 15 days | 0.0372 | 0.1055 |
| | | 2 | 15 or more days | -0.4327 | 0.1653 |
| A3, A4 | Time Parents spend listening to radio per week | 0 | Less than 1 hour | 0.2697 | 0.0828 |
| | | 1 | A few hours | 0.0333 | 0.0708 |
| | | 2 | Several hours | -0.3030 | 0.0828 |
| A6 | Parent reading of magazines | 0 | Never | 0.3383 | 0.138 |
| | | 1 | Sometimes | -0.0823 | 0.0882 |
| | | 2 | often | -0.2559 | 0.0914 |
| H5 | Parental martial status | 0 | Not married | 0.1663 | 0.0631 |
| | | 1 | Married | -0.1663 | 0.0631 |
| A5 | Parental reading of newspapers | 0 | Never | -0.0462 | 0.1482 |
| | | 1 | Sometimes | 0.2250 | 0.1079 |
| | | 2 | Often | -0.1788 | 0.0956 |
| A1, A2 | Time parents spend watching TV per week | 0 | Less than 1 hour | 0.4477 | 0.1437 |
| | | 1 | A few hours | -0.0547 | 0.0889 |
| | | 2 | Several hours | -0.3930 | 0.0935 |
| A3, A4, A5 | Youth was White | 0 | No | -0.1695 | 0.0718 |
| | | 1 | Yes | 0.1695 | 0.0718 |

**Table C-I. Stable Model for parent specific exposure index among all parents of youth aged 9 to 18**

| Quex Items | Description of variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of households with income above $75,000/year in the neighborhood in 1990 | Continuous(0-100) | | 0.0248 | 0.00462 |
| | Percentage of housing in the neighborhood in 1990 that was large structures with 50 or more housing units. | Continuous(0-100) | | 0.0209 | 0.00594 |
| | Talking belief/Attitude index | Continuous | | -0.00127 | 0.000612 |
| A9a, A9b | Parent viewing of BET and Spanish-language | 0 | None | 0.4265 | 0.0844 |
| | Cable channels in the past 30 days | 1 | Less than 15 days | 0.1246 | 0.0934 |
| | | 2 | 15 or more days | -0.5510 | 0.1305 |
| A3, A4 | Time Parents spend listening to radio per week | 0 | Less than 1 hour | 0.3864 | 0.0812 |
| | | 1 | A few hours | -0.0677 | 0.0682 |
| | | 2 | Several hours | -0.3187 | 0.0785 |
| A1, A2 | Time parents spend watching TV per week | -1 | Less than 1 hour | 0.7392 | 0.1425 |
| | | 0 | A few hours | -0.1494 | 0.0875 |
| | | 1 | Several hours | -0.5898 | 0.0917 |
| | Gender of parents | 1 | Male | 0.1263 | 0.0551 |
| | | 2 | Female | -0.1263 | 0.0551 |
| B2, B3b, B3c | Parent talking behavior | 0 | Never | 0.1578 | 0.1285 |
| | | 1 | Seldom | 0.0002 | 0.1309 |
| | | 2 | Sometimes | 0.1285 | 0.0982 |
| | | 3 | Often | -0.2865 | 0.0876 |

**Table C-J. Lagged Model for youth general exposure index among youth aged 12 to 18 at Wave 4 who had never tried marijuana at Wave 1**

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of persons in the neighborhood in 1990 who are urban and live inside urbanized areas | Continuous (0-100) | | 0.00253 | 0.00101 |
| | Percentage of persons who are American Indian, Eskimo or Aleut in the neighborhood in 1990 | Continuous (0-100) | | -0.0264 | 0.0114 |
| A27, A28, A28a, A28b | Youth consumption of specific cable channels targeted by the Media Campaign | Continuous range from 0 (did not watch) to 3 (watched several channels often) | | 0.2503 | 0.0883 |
| C34a-d | Youth likes new and exciting friends and experiences | Continuous range from –2 (dislikes) to +2 (strongly likes) | | -0.3278 | 0.1364 |
| C35a, C35b | Youth perception of parental awareness of youth activities and plans | Continuous range from –2 (never aware) to +2 (always aware) | | 0.2672 | 0.0439 |
| C31a – C31d | Youth association with antisocial peers | Continuous range from 0 (did not associate) to 6 (associated often) | | 0.1393 | 0.0605 |
| A19 | Time youth spends watching TV on an average weekday | -1 | Less than 1 hour | -0.3723 | 0.0745 |
| | | 0 | A few hours | 0.0977 | 0.0639 |
| | | 1 | Several hours | 0.2746 | 0.0870 |
| A24 | Youth reading of magazines | 0 | Rarely | -0.1614 | 0.0509 |
| | | 1 | Often | 0.1614 | 0.0509 |
| C30 | Time spent with friends without adult supervision | 0 | Never or some | 0.1846 | 0.0631 |
| | | 1 | Often or always | -0.1846 | 0.0631 |
| D27 | Youth use of the Internet | -1 | Rarely | -0.2205 | 0.0701 |
| | | 0 | Sometimes | -0.0601 | 0.0623 |
| | | 1 | Every day | 0.2805 | 0.0735 |
| C34c, C34d | Youth score on sensation seeking tendencies (median split) | 0 | Low | 0.1062 | 0.0276 |
| | | 1 | High | -0.1062 | 0.0276 |
| A9a, A9b | Parent viewing of BET and Spanish-language cable channels in the past 30 days | 0 | None | -0.2610 | 0.0688 |
| | | 1 | Less than 15 days | 0.0080 | 0.0630 |
| | | 2 | 15 or more days | 0.2530 | 0.0742 |
| H5 | Parental martial status | 0 | Not Separated | -0.2583 | 0.1142 |
| | | 1 | Separated | 0.2583 | 0.1142 |
| A3, A4, A5 | Youth was Black | 0 | No | -0.5251 | 0.1595 |
| | | 1 | Yes | 0.5251 | 0.1595 |
| | Youth was aged between 16 and 18 | 0 | No | 0.2191 | 0.0839 |
| | | 1 | Yes | -0.2191 | 0.0839 |

### C.4.4.10 Lagged Model for the Youth Recall-Aided Exposure Index

The lagged model for the youth recall-aided exposure index found 14 significant variables using stepwise regression. However, tests of balance indicted that the model needed to be improved. Therefore, several variables plus age interactions were added to the model. The resulting model has 46 main terms plus age interactions. These variables are presented, together with their coefficients, in Table C-K.

### C.4.4.11 Lagged Model for the Parent General Exposure Index

There were 11 significant variables in the lagged model for parental general exposure. These and their coefficients are tabulated in Table C-L.

### C.4.4.12 Lagged Model for the Parent Recall-Aided Exposure Index

The lagged model for parental recall-aided exposure found nine significant variables. These variables and their coefficients are presented in Table C-M.

## C.5 Testing for Significance of Counterfactual Effects

Several approaches were employed to assess the significance of estimated effects. The actual mean on each outcome for the weighted sample and all of the counterfactual means for each exposure group were displayed with their confidence intervals and were available for visual inspection. The population effect (called the "direct effect") was assessed by comparing the actual mean with the counterfactual mean for the lowest exposure group. This was done by estimating the variance of the direct effect and using that to place a confidence interval on the direct effect. The second approach was to estimate the variance of the maximum effect, the difference between the lowest and highest exposure groups, and use that to place a confidence interval on the maximum effect. The third was to adapt a test (the Jonckheere-Terpstra) for monotone dose-response relationship. The monotone dose-response test assessed the overall association between exposure and outcome, whereas the direct effect test is an estimate of the average effect in the population, while the maximum effect test provides a hypothetical estimate of the effect if all respondents received the highest does. With all of the approaches, the extra variance introduced by complex sample design, nonresponse adjustment, and counterfactual projection were reflected as fully as possible.

### C.5.1 Estimating Variances on Counterfactual Projections

Replicate weights had been prepared for variance estimation of ordinary survey statistics as explained in Appendix A. There are 100 of these replicate weights for every subject. The process of adjusting the standard survey weights for counterfactual projection was partially repeated on each set of replicate weights. As explained in Section C.4.1 of this appendix, there were four major steps in this process. The first was to model exposure. The second was to create a partition of the data set based on the values of $X_i\hat{\beta}$. The third was to estimate the exposure propensity within each cell of the partition for each of the different exposure levels. The fourth was to apply the inverse of these estimated propensities to the sampling weights. To estimate the variances of the counterfactual projections, only the third and fourth steps were replicated. The first two were not. Ideally, all the steps would have

been replicated, but technical issues made this infeasible. As a result, the variance estimates are likely to be a little too small and the confidence intervals a little tighter than they should be.

**Table C-K. Lagged Model for youth specific exposure index among youth aged 12 to 18 at Wave 4 who had never tried marijuana at Wave 1**

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Youth's Age at Wave 4 | Ordinal values from 12 to 18 | | 0.0416 | 0.0871 |
| A12a | Youth school was in session in last 30 days | 0 | Yes | -1.3465 | 0.8958 |
| | | 1 | No | 1.3465 | 0.8958 |
| A19 | Time youth spends watching TV on an average weekday | -1 | Less than 1 hour | -0.3624 | 0.1107 |
| | | 0 | A few hours | 0.1614 | 0.0857 |
| | | 1 | Several hours | 0.2010 | |
| A20 | Time youth spends watching TV on weekends | -1 | Less than 2 hours | -0.0595 | 0.1163 |
| | | 0 | Several hours | 0.0456 | 0.0824 |
| | | 1 | More than 9 hours | 0.0139 | |
| A22 | Time youth spends listening to the radio on weekends | -1 | Less than 2 hours | 0.0771 | 0.0973 |
| | | 0 | Several hours | 0.0936 | 0.0979 |
| | | 1 | More than 9 hours | -0.1707 | |
| A22b | Youth language of radio listening | 0 | Spanish | -0.5582 | 0.3250 |
| | | 1 | English | 0.5582 | 0.3250 |
| A24 | Youth reading of magazines | 0 | Rarely | -0.0319 | 0.0742 |
| | | 1 | Often | 0.0319 | 0.0742 |
| A27, A28, A28a, A28b | Youth consumption of specific cable channels targeted by the Media Campaign | Continuous range from 0 (did not watch) to 3 (watched several channels often) | | 0.3232 | 0.1158 |
| C30 | Time spent with friends without adult supervision | 0 | Never or some | 0.3489 | 0.1192 |
| | | 1 | Often or always | -0.3489 | 0.1192 |
| C31a – C31d | Youth association with antisocial peers | Continuous range from 0 (did not associate) to 6 (associated often) | | -0.0129 | 0.0711 |
| C34a-d | Youth likes new and exciting friends and experiences | Continuous range from –2 (dislikes) to +2 (strongly likes) | | -0.0445 | 0.0901 |
| C34aa-ac | Youth behaving anti-socially in last 12 months | Continuous range from –2 (never) to +2 (5 or more times) | | -0.0395 | 0.1690 |
| C35a, C35b | Youth perception of parental awareness of youth activities and plans | Continuous range from –2 (never aware) to +2 (always aware) | | -0.0105 | 0.0620 |
| C36c | Youth fought or argued with a parent in the last 30 days | 0 | Never or sometimes | -0.2129 | 0.0793 |
| | | 1 | Often or always | 0.2129 | 0.0793 |
| D27 | Youth use of the Internet | -1 | Rarely | -0.2027 | 0.0968 |
| | | 0 | Sometimes | 0.0013 | 0.0836 |
| | | 1 | Every day | 0.2014 | |
| C5c (parents) | How often parents did activities with children in past week | -3 | Not at all | -0.2750 | 0.1410 |
| | | -2 | Once | 0.0208 | 0.1414 |
| | | -1 | Twice | 0.0660 | 0.1335 |
| | | 0 | 3 times | -0.1532 | 0.1511 |
| | | 1 | 4 or 5 times | 0.0884 | 0.1514 |
| | | 2 | 6 or 7 times | 0.1718 | 0.2317 |
| | | 3 | More than 7 times | 0.0812 | |

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| C10c (parents) | Parental perception of fighting with the youth | -2 | Never | 0.1443 | 0.1290 |
| | | -1 | Sometimes | -0.0046 | 0.1143 |
| | | 0 | Half the time | 0.2379 | 0.1723 |
| | | 1 | Often | -0.5110 | 0.2219 |
| | | 2 | Always | 0.1334 | |
| | Percentage of persons in the neighborhood in 1990 who were urban and lived inside urbanized areas | | | 0.0016 | 0.0014 |
| | Percentage of persons in the neighborhood in 1990 who lived on farms | | | -0.0221 | 0.0144 |
| | Percentage of persons in the neighborhood in 1990 who were Asian or Pacific Islander | | | -0.0061 | 0.0091 |
| | Percentage of persons in the neighborhood in 1990 who were Hispanic | | | -0.0039 | 0.0081 |
| | Percentage of persons in the neighborhood in 1990 who were Mexican American | | | -0.0064 | 0.0122 |
| | Percentage of persons in the neighborhood in 1990 who were Puerto Rican | | | 0.0056 | 0.0146 |
| | Percentage of persons in the neighborhood in 1990 who were Cuban American | | | 0.0004 | 0.0162 |
| | Percentage of persons in the neighborhood in 1990 who were institutionalized | | | 0.0481 | 0.0232 |
| | Percentage of persons in the neighborhood in 1990 who lived in noninstitutional group quarters | | | 0.0369 | 0.1684 |
| | Percentage of persons in the neighborhood in 1990 who had at least a college degree | | | -0.0102 | 0.0053 |
| | Percentage of persons in the neighborhood in 1990 16 and older and employed in mining | | | -0.0624 | 0.0571 |
| | Percentage of persons in the neighborhood in 1990 16 and older with farming, forestry or fishing occupations | | | 0.0125 | 0.0205 |
| | Neighborhood is classified as rural | 0 | No | 0.0275 | 0.0821 |
| | | 1 | Yes | -0.0275 | 0.0821 |
| C34a | Youth score on sensation seeking tendencies (median split) | 0 | Low | 0.0631 | 0.0670 |
| | | 1 | High | -0.0631 | 0.0670 |
| A8, A9 | Youth current or last school grade | -1 | Primary (Grades 1-6) | -0.0204 | 0.1591 |
| | | 0 | Middle (Grades 7-9) | 0.1234 | 0.1058 |
| | | 1 | High (Grades 10-12) | -0.1030 | |
| B6 - B7 | Youth alcohol consumption | 0 | Low | 0.0496 | 0.0762 |
| | | 1 | High | -0.0496 | 0.0762 |
| A9 (parents) | Household has cable or satellite TV service | 0 | No | 0.1402 | 0.0618 |
| | | 1 | Yes | -0.1402 | 0.0618 |

| Quex Items | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| H3 (parents) | Parent report of youth's average grades | 0 | Low | -0.1169 | 0.0617 |
| | | 1 | High | 0.1169 | 0.0617 |
| H1, H2 (parents) | Child's school grade span | -1 | Elem. or Middle School | 0.1618 | 0.1008 |
| | | 0 | Middle + High | 0.0727 | 0.1035 |
| | | 1 | High school only | -0.2345 | |
| F9 (parents) | Parental use of the internet | -1 | Rarely | -0.0650 | 0.0716 |
| | | 0 | Sometimes | -0.1354 | 0.0860 |
| | | 1 | Every day | 0.2004 | |
| A5 (parents) | Parental reading of newspapers | -1 | Never | -0.0283 | 0.1249 |
| | | 0 | Sometimes | -0.1390 | 0.0940 |
| | | 1 | Often | 0.1673 | |
| H6 (parents) | Parental educational attainment | 0 | HS or less | -0.0599 | 0.0553 |
| | | 1 | Beyond HS | 0.0599 | 0.0553 |
| | Parent has child aged 9 to 11 | 0 | No | -0.0917 | 0.1332 |
| | | 1 | Yes | 0.0917 | 0.1332 |
| | Parent has child aged 14 to 18 | 0 | No | 0.4800 | 0.6362 |
| | | 1 | Yes | -0.4800 | 0.6362 |
| H5 (parents) | Parent is married | 0 | No | 0.0876 | 0.0738 |
| | | 1 | Yes | -0.0876 | 0.0738 |
| A2-A5 | Youth's race is white | 0 | No | 0.1899 | 0.0882 |
| | | 1 | Yes | -0.1899 | 0.0882 |
| A2-A5 | Youth's race/ethnicity is Hispanic | 0 | No | -0.0942 | 0.1317 |
| | | 1 | Yes | 0.0942 | 0.1317 |
| A16 | Youth plans to graduate from a two-year college program | 0 | No | 0.0345 | 0.1240 |
| | | 1 | Yes | -0.0345 | 0.1240 |
| | Youth gender | 0 | Male | -0.2036 | 0.0656 |
| | | 1 | Female | 0.2036 | 0.0656 |
| A13b, A13cb | Youth report of cut school days | -1.5 | Low | -2.6349 | 157.0000 |
| | | -0.5 | Med-Low | -3.0756 | 157.0000 |
| | | 0.5 | Med-High | -4.0931 | 157.0000 |
| | | 1.5 | High | 9.8036 | |
| G6 (parents) | Parent has ever used marijuana | 0 | No | -0.0488 | 0.0627 |
| | | 1 | Yes | 0.0488 | 0.0627 |

**Table C-L. Lagged Model for parent general exposure index among all parents of youth aged 12 to 18**

| Quex Items | Description of variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of persons who are urban and live inside urbanized areas in the neighborhood in 1990 | Continuous(0-100) | | 0.00359 | 0.0012 |
| | Percentage of households with income above $75,000/year in the neighborhood in 1990 | Continuous(0-100) | | -0.0216 | 0.00502 |
| | Percentage of single-family structures | Continuous(0-100) | | 0.00477 | 0.00231 |
| | Talking belief/Attitude index | Continuous | | 0.00133 | 0.000613 |
| A9a, A9b | Parent viewing of BET and Spanish-language cable channels in the past 30 days | 0 | None | -0.5808 | 0.0945 |
| | | 1 | Less than 15 days | -0.1854 | 0.1016 |
| | | 2 | 15 or more days | 0.7662 | 0.1524 |
| A3, A4 | Time Parents spend listening to radio per week | 0 | Less than 1 hour | -0.1661 | 0.0818 |
| | | 1 | A few hours | -0.0316 | 0.0691 |
| | | 2 | Several hours | 0.1977 | 0.0800 |
| A6 | Parent reading of magazines | 0 | Never | -0.3902 | 0.1345 |
| | | 1 | Sometimes | 0.0363 | 0.0859 |
| | | 2 | often | 0.3539 | 0.0892 |
| H5 | Parental martial status | 0 | Not married | -0.176 | 0.0589 |
| | | 1 | Married | 0.176 | 0.0589 |
| A5 | Parental reading of newspapers | 0 | Never | -0.1389 | 0.1432 |
| | | 1 | Sometimes | -0.1089 | 0.1051 |
| | | 2 | Often | 0.2478 | 0.0928 |
| A1, A2 | Time parents spend watching TV per week | 0 | Less than 1 hour | -0.417 | 0.1439 |
| | | 1 | A few hours | 0.1217 | 0.0885 |
| | | 2 | Several hours | 0.2952 | 0.0926 |
| B2, B3b, B3c | Parent talking behavior | 0 | Never | -0.4303 | 0.1309 |
| | | 1 | Seldom | 0.2637 | 0.1339 |
| | | 2 | Sometimes | -0.0154 | 0.099 |
| | | 3 | Often | 0.1820 | 0.0886 |

**Table C-M. Lagged Model for parent specific exposure index among all parents of youth aged 12 to 18 at Wave 4**

| Quex Items | Description of variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| | Percentage of households with income above $75,000/year in the neighborhood in 1990 | Continuous(0-100) | | -0.0144 | 0.00446 |
| | Percentage of housing in the neighborhood in 1990 that was large structures with 50 or more housing units. | Continuous(0-100) | | -0.0179 | 0.00581 |
| A3, A4 | Time Parents spend listening to radio per week | -1 | Less than 1 hour | -0.189 | 0.0785 |
| | | 0 | A few hours | 0.0041 | 0.0655 |
| | | 1 | Several hours | 0.1849 | 0.0747 |
| A2b | Language of parental TV viewing | 0 | Spanish | 0.3964 | 0.1350 |
| | | 1 | English | -0.3964 | 0.1350 |
| A1, A2 | Time parents spend watching TV per week | -1 | Less than 1 hour | -0.3837 | 0.1393 |
| | | 0 | A few hours | 0.1297 | 0.0854 |
| | | 1 | Several hours | 0.2540 | 0.0880 |
| | Respondent is Hispanic | 0 | No | -0.2613 | 0.0978 |
| | | 1 | Yes | 0.2613 | 0.0978 |
| G6 | Parent has ever used marijuana | 0 | No | 0.1136 | 0.0504 |
| | | 1 | Yes | -0.1136 | 0.0504 |
| B2, B3b, B3c | Parent talking behavior | 0 | Never | -0.1721 | 0.1224 |
| | | 1 | Seldom | -0.0418 | 0.1258 |
| | | 2 | Sometimes | -0.1116 | 0.0952 |
| | | 3 | Often | 0.3254 | 0.0778 |

The reason for this is that confidence intervals do not reflect the uncertainty due to selecting the most important predictors of exposure. Different samples would no doubt have resulted in different choices of which variables to include in the ordinal logit model. However, the extra uncertainty introduced by model selection among the variables considered is probably small. Note that the confidence intervals are also conditioned on the assumptions made about exposure. If there were important covariates that were omitted from the modeling process because they were never asked in the questionnaire, the confidence intervals will not provide the 95 percent coverage promised.

Let $w_{itr}$ be the *r*-th replicated counterfactual weight for the *t*-th exposure level for the *i*-th observation. Let $w_{it0}$ be the full sample counterfactual weight. Note that these weights are equal to zero for the *i*-th observation unless the *i*-th observation actually experienced the *t*-th exposure level. Let $\delta_{it}$ be an indicator flag for the *t*-th exposure level for the *i*-th observation. A unified set of counterfactual weights was then created by stacking these weights according to

$$w'_{ir} = \sum_k \delta_{ik} w_{ikr} \text{ and } w'_{i0} = \sum_k \delta_{ik} w_{ik0} .$$

The counterfactual mean for some outcome *y* on some class *c* indicated by $\varepsilon_{ci}$ and exposure level *t* is then

$$\hat{y}_{ct} = \frac{\sum_i w'_{i0}\delta_{it}\varepsilon_{ci}y_i}{\sum_i w'_{i0}\delta_{it}\varepsilon_{ci}} \text{ with variance estimate } \operatorname{var}\hat{y}_{ct} = \sum_r b_r \left( \frac{\sum_i w'_{ir}\delta_{it}\varepsilon_{ci}y_i}{\sum_i w'_{ir}\delta_{it}\varepsilon_{ci}} - \frac{\sum_i w'_{i0}\delta_{it}\varepsilon_{ci}y_i}{\sum_i w'_{i0}\delta_{it}\varepsilon_{ci}} \right)^2 ,$$

where the $b_r$ are factors chosen to correspond to the replication method.

## C.5.2 Confidence Intervals on Direct Effects

The direct effect is defined as the difference between the actual estimate and the counterfactual estimate for the low exposure category. To estimate the variance on this effect, the first step was to estimate the covariance between a counterfactual estimate and an actual estimate as

$$\operatorname{covar}(\hat{y}_{ct}, \hat{y}_c) = \sum_r b_r \left( \frac{\sum_i w'_{ir}\delta_{it}\varepsilon_{ci}y_i}{\sum_i w'_{ir}\delta_{it}\varepsilon_{ci}} - \frac{\sum_i w'_{i0}\delta_{it}\varepsilon_{ci}y_i}{\sum_i w'_{i0}\delta_{it}\varepsilon_{ci}} \right) \left( \frac{\sum_i w_{ir}\delta_{it}\varepsilon_{ci}y_i}{\sum_i w_{ir}\delta_{it}\varepsilon_{ci}} - \frac{\sum_i w_{i0}\delta_{it}\varepsilon_{ci}y_i}{\sum_i w_{i0}\delta_{it}\varepsilon_{ci}} \right) .$$

In the second step, the variance on the direct effect was estimated as

$$\operatorname{var}(\hat{y}_c - \hat{y}_{ct}) = \operatorname{var}(\hat{y}_c) + \operatorname{var}(\hat{y}_{ct}) - 2\operatorname{covar}(\hat{y}_{c,}\hat{y}_{ct}), \text{ where } t = 1.$$

Confidence intervals on maximum effects are calculated using WESVAR and a description can be found in Appendix A, section A.3.1.

## C.5.4 Testing for a Monotone Dose-Response Relationship

A standard nonparametric test in toxicology and biopharmaceutical research for does-response relationship is the Jonckheere-Terpstra test. This test is described in the SAS manual among other

places. It is appropriate for testing whether two ordinal variables have a monotone relationship to each other. It does not require that the response (outcome) variable have a normal distribution, as is the case in standard analysis of variance procedures. This is important in this report because the outcomes of interest are generally not normally distributed. In this application, a monotone relationship is a relationship such that as the level of exposure increases, the level of the outcome variable moves in one direction only. There is no requirement that the outcome rise linearly or steadily. It can rise in jerks and pauses, but there can be no reversals. In terms of the cognitive processes, it is assumed that extra exposure to advertising will either have an effect or not have an effect, but that the direction of the effect will never reverse. Although it might be possible to imagine a situation where light exposure is beneficial while heavy exposure actually has the opposite of the desired effect, this does not seem plausible in general.

Prior applications of the Jonckheere-Terpstra (JT) test were made only to simple random samples. In this, the counterfactual weights carry the information about confounders needed to remove their effect from the association. Furthermore, the survey clustering introduces correlations between observations that violate the standard JT assumption of independent observations. The test was therefore modified for this application.

SAS has an option to use a weight in calculating the JT test. This feature was used. If a subject has a weight of *W*, using the weight has the same effect on the calculations as if *W* copies of the subject were included in the database. Since the weights were in the tens of thousands, SAS perceives the sample size as being much larger than it really is and returns inappropriate significance levels. This was corrected by replicating the JT.

Let $J_0$ be the value of the JT test Z-statistic produced by SAS using the full sample counterfactual weights $w'_{i0}$ and $J_r$ be the value of the JT test produced by SAS using the *r*-th replicated counterfactual weights $w'_{ir}$. The variance on the JT statistic was calculated as

$$v = \sum_{r=1}^{100} b_r (J_r - J_o)^2 \ .$$

The corrected JT test is then given as

$$J_C = \frac{J_0}{\sqrt{v}} \ .$$

Under the null hypothesis that there is no relationship between exposure and the outcome, the statistic $J_C$ has an approximate t-distribution with 100 degrees of freedom. So the alternate hypothesis of a monotone relationship between exposure and outcome is accepted if $J_C$ >1.98.

# References

D'Agostino, R.B., Jr. (1998). Tutorial in biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine,* 17, 2265-2281.

Drake, C. (1993). Effects of misspecification of the propensity score on estimators of treatment effect. Vol. 49, 1231-36. Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, 87, 706-710.

Joffe, M. M., and Rosenbaum, P. R. (1990). Propensity scores. *American Journal of Epidemiology*, 150, 327-333.

Rosenbaum, P.R., and Rubin, D.B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association,* 79, 516-524.

Rosenbaum, P.R., and Rubin, D.B. (1983). The central role of propensity score in observational studies for causal effects. *Biometrika,* 70, 41-55.

Imbens, G.W., (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*. 87. 706-710.

# Appendix D

# Waves 1, 2, 3 and 4 – NSPY Anti–drug Advertisements Shown to Respondents

## Wave 4

Table D-1a. Television Advertisements Shown to Parents

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Derrick Brooks | NFL player Derrick Brooks talks about how parents can keep kids drug free by making time and monitoring them. |
| | Eddie George | NFL player Eddie George talks about how his mother kept him from using drugs. |
| | My Hero GM | A series of young boys and girls address their parents about the necessary actions they need to take to teach them about the dangers of drugs. Parents need to be the grown-up. |
| African American | Thanks | A series of teens thank their parents for monitoring, disciplining and being there for them during troublesome times. |
| | Derrick Brooks | NFL player Derrick Brooks talks about how parents can keep kids drug free by making time and monitoring them. |
| | Eddie George | NFL player Eddie George talks about how his mother kept him from using drugs. |
| | My Hero AA | A series of African American young boys and girls address their parents about the necessary actions they need to take to teach them about the dangers of drugs. Parents need to be the grown-up. |
| | Thanks | A series of teens thank their parents for monitoring, disciplining and being there for them during troublesome times. |
| Hispanic | Alert (Spanish) | Ad warns parents about the dangers of everyday products that can be used to get high by teens. Parents are encouraged to be curious about what their children are doing, even when there is no reason to suspect drug use. |
| | Shadow – Brochure (Spanish) | A Hispanic boy is “shadowed” by the presence of drugs in society. His concerned parents turn to the brochure they got about drug abuse for advice about talking to the son. |
| | Shadow – Monitoring (Spanish) | A Hispanic girl is “shadowed” by the presence of drugs in society. Her concerned father realizes the importance of monitoring his daughter’s activities and friends. |

# Wave 4 (continued)

## Table D-1b. Radio Advertisements Played for Parents

| Target Audience | Ad name | Description |
|---|---|---|
| General market | My Hero GM | A series of young boys and girls address their parents about the necessary actions they take to teach them about the dangers of drugs. Parents need to be the grown-up. |
| | Sooner or Later David | Teen is being lectured by parent about the dangers of taking and sharing ecstasy with friends, especially when purchased from a stranger. Talk to youth "sooner" rather than "later." |
| | Sooner or Later Megan | An angry parent is on the phone with her incoherent daughter after learning that she used ecstasy. Message is for parents to speak with youth "sooner" rather than "later." |
| | Thanks | A series of teens thank their parents for disciplining and being there for them during troublesome times. |
| African American | My Hero AA | A series of young African American boys and girls address their parents about the necessary actions they take to teach them about the dangers of drugs. Parents need to be the grown-up. |
| | Thanks | A series of teens thank their parents for disciplining and being there for them during troublesome times. |
| Hispanic | Alert –Dad (Spanish) | Hispanic male warns parents about the dangers of everyday products that can be used to get high by teens. Parents are encouraged to be curious about what their children are doing, even when there is no reason to suspect drug use. |
| | Alert-Mom (Spanish) | Hispanic female warns parents about the dangers of everyday products that can be used to get high by teens. Parents are encouraged to be curious about what their children are doing, even when there is no reason to suspect drug use. |
| | Shadow – Monitoring (Spanish) | A Hispanic girl is "shadowed" by the presence of drugs in society. Her concerned father realizes the importance of monitoring his daughter's activities and friends. |

# Wave 4 (continued)

## Table D-1c. Television Advertisements Shown to Youth

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Being Myself | Animation of young girl in various activities: cheerleading, playing basketball, studying. When offered drugs, she blows them off. Her future is her anti-drug. |
| | Brain | Graphical depiction of a person's head when using inhalants. Be nice to your brain – don't use inhalants. |
| | Brothers | Younger brother is shown shadowing his older brother, wanting to emulate him. Older brother is offered a joint, younger brother watches to see what he'll do. |
| | Derrick Brooks | NFL player Derrick Brooks talks about having self-respect and not using drugs. |
| | Drawing | Sketch work shows a young artist transforming drug users into foolish characters and nonusers into popular winners. Drawing is the youth's anti-drug. |
| | Music/Mix Tapes | Animation of youth walking around city streets, listening to music. Youth encounters negative drug influences but continues listening to the music. Youth states that music is his anti-drug. |
| | Tiki Barber | NFL player Tiki Barber talks about how drugs can keep you from achieving your goals. Football is his anti-drug. |
| African American | Vision Warrior | Young man talks about how smoking marijuana led him to use harder drugs. |
| | Derrick Brooks | NFL player Derrick Brooks talks about having self-respect and not using drugs. |
| | Music/Mix Tapes | Animation of youth walking around city streets, listening to music. Youth encounters negative drug influences but continues listening to the music. Youth states that music is his anti-drug. |
| | Tiki Barber | Tiki Barber of the NY Giants talks about how drugs can keep you from achieving your goals. Football is his anti-drug. |
| Hispanic | Drowning (Spanish) | Young girl is shown as drowning in her own room, unable to escape. This is the way your brain feels when you use inhalants. |
| | La Musica (Spanish) | Animation of youth walking around city streets, listening to music. Youth encounters negative drug influences but continues listening to the music. Youth states that music is his anti-drug. |

# Wave 4 (continued)

## Table D-1d. Radio Advertisements Played for Youth

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Basketball | Young male explains why basketball is his anti-drug. |
| | (Two) Brothers | Younger brother brags about his older brother's accomplishments. When the older brother is offered drugs, he realizes he sets the example for his younger brother. |
| | Cross Country | Young male explains why cross-country running is his anti-drug. |
| | Excuses | Excuses you can give for not smoking marijuana are provided. |
| | Limericks | Young male recites limerick about dangers of drug use – writing limericks is his anti-drug. |
| | Margot | Female youth has a younger friend with a disability and wants to be her role model. Teaching her about life is more important than taking drugs. Her younger friend is her anti-drug. |
| | The Rant | Ad talks about the lies associated with ecstasy when viewed by nonusers. |
| African American | Basketball | Young male explains why basketball is his anti-drug. |
| Hispanic | Jose (Spanish) | Jose is a teen whose anti-drug is music. He sings part of a song called "La Rosa" in the ad. |
| | She Did It (Spanish) | Girls talk to popular girl who says no to marijuana and is still popular. |

# Wave 3

## Table D-1a. Television Advertisements Shown to Parents

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Clinic | A father and son are shown walking through a clinic – like setting, but finally arrive at a basketball clinic. The ad offers a telephone number to get a book on parent – child activities. |
| | My Hero | A series of young boys and girls address their parents about the necessary actions they need to take to teach them about the dangers of drugs. Parents need to be the grown-up. |
| | Needle/Spray Can | Ad relays message to parents about unsuspecting drugs under the sink in the home. Aerosol can is depicted as a syringe. Inhalants are dangerous and deadly. "Communication" is the anti-drug. |
| | Smoke | Ad opens with two smoke streams and a verbal message about parental interaction with kids. During message, the camera follows the smoke streams to two roasting marshmallows over a campfire. Parents are the anti-drug. |
| | Thanks | A series of teens thank their parents for disciplining and being there for them during troublesome times. |
| African American | Clinic | A father and son are shown walking through a clinic – like setting, but finally arrive at a basketball clinic. The ad offers a telephone number to get a book on parent – child activities. |
| | Deal | Father is imitating a drug dealer to his son on a playground to see how he reacts. The boy refuses the offer in a stern fashion to his father's delight. |
| | My Hero | A series of African American young boys and girls address their parents about the necessary actions they need to take to teach them about the dangers of drugs. Parents need to be the grown-up. |
| Hispanic | Mirrors – (Spanish) | A boy wanders through a house of mirrors while his parents search for him. "Your child can be under the illusion that smoking marijuana is harmless." It isn't. |
| | Needle/Spray Can (Spanish) | Ad relays message to parents about unsuspecting drugs under the sink in the home. Aerosol can is depicted as a syringe. Inhalants are dangerous and deadly. "Communication" is the anti-drug. |
| | Shadow – Brochure (Spanish) | A Hispanic boy is "shadowed" by the presence of drugs in society. His concerned parents turn to the brochure they got about drug abuse for advice about talking to the son. |

# Wave 3 (continued)

## Table D-1b. Radio Advertisements Played for Parents

| Target Audience | Ad name | Description |
|---|---|---|
| General market | Basketball | Activities are listed that kids would rather do than drugs. The number one deterrent to drugs is parents and the time spent with their kids. |
| | Desperate | Ad opens with what sounds like a parent lecturing the son about the dangers of drugs. However, the parent is actually playing a video game with the youth and spending time with him. Phone number and web site is given for information about keeping youths off drugs. |
| | Happy Birthday Steven | A mother describes what she does (feeding, bathing) to take care of her teenaged son who used inhalants and suffered brain damage. |
| | Kathy Abel | A woman describes how her son died from sniffing fumes with his friends. Youths and adults need to be informed about the lethal dangers with the seemingly "harmless" fun of inhalant use. |
| | Keep Trying | A boy describes all the times he was told by his parent to keep trying. He encourages parents to "keep trying" to talk to kids about marijuana. |
| | My Hero | A series of young boys and girls address their parents about the necessary actions they take to teach them about the dangers of drugs. Parents need to be the grown-up. |
| | Needle/Spray Can | Message informs parents about the dangers of inhalants in the home. Phone number and web site is given for more information. Communication is the anti-drug. |
| | Sooner or Later David | Teen is being lectured by parent about the dangers of taking and sharing ecstasy with friends, especially when purchased from a stranger. Talk to youth "sooner" rather than "later." |
| | Sooner or Later Megan | An angry parent is on the phone with her incoherent daughter after learning that she used ecstasy. Message is for parents to speak with youth "sooner" rather than "later." |
| | Symptoms | Ad talks about the negative ripple effects that occur in the family when a member is using marijuana. Examples include depression, withdrawal, and hostility. |
| | Tree Fort | Activities are suggested to do with your kids: rollerblade, play chess, go to movie. Be aware of at-risk hours—between 4 pm and 6 pm is when kids are most likely to try drugs. |
| African American | Keep Trying | A boy describes all the times he was told by his parent to keep trying. He encourages parents to "keep trying" to talk to kids about marijuana. |
| | My Hero | A series of young African American boys and girls address their parents about the necessary actions they take to teach them about the dangers of drugs. Parents need to be the grown-up. |

# Wave 3 (continued)

## Table D-1b. Radio Advertisements Played for Parents

| Target Audience | Ad name | Description |
|---|---|---|
| Hispanic | Happy Birthday Raoul (Spanish) | A mother describes what she does (feeding, bathing) to take care of her teenaged son who used inhalants and suffered brain damage. |
| | Needle/Spray Can (Spanish) | Message informs parents about the dangers of inhalants in the home. Phone number and web site is given for more information. Communication is the anti-drug. |
| | Pepperoni (Spanish) | The best way to keep youth younger than 15 from using drugs is by supervising them and being an effective parent. |
| | Shadow – Brochure (Spanish) | A Hispanic boy is "shadowed" by the presence of drugs in society. His concerned parents turn to the brochure they got about drug abuse for advice about talking to their son. |

# Wave 3 (continued)

## Table D-1c. Television Advertisements Shown to Youth

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Dance | Animation of a girl dancing to music on her radio. While dancing, she is offered drugs by two boys. She refuses the offer and states that dancing is her anti-drug. |
| | DJ | A boy talks about his feelings when he performs as a disk jockey. Asks "what's your anti-drug?" |
| | Drawing | Sketch work shows an young artist transforming drug users into foolish characters and nonusers into popular winners. Drawing is the youth's anti-drug. |
| | Football | A football player talks about catching a pass. Asks "what's your anti-drug?" |
| | Friends | A boy talks about doing everything with his friends and sticking together with them. Asks "what's your anti-drug?" |
| | Icon | Ad shows a collage of images of various activities. Asks "what's your anti-drug?" |
| | It's OK to Pass | Group of suburban youths sit in a garage talking and passing a drug to each other. The last youth rejects the drug and passes it on. Her rejection is acceptable to her peers indicating that it's 'ok' to pass. |
| | Music/Mix Tapes | Animation of youth walking around city streets, listening to music. Youth encounters negative drug influences but continues listening to the music. Youth states that music is his anti-drug. |
| | Swimming | A girl talks about how much she enjoys swimming. Asks "what's your anti-drug?" |
| African American | DJ | A boy talks about his feelings when he performs as a disk jockey. Asks "what's your anti-drug?" |
| | Football | A football player talks about catching a pass. Asks "what's your anti-drug?" |
| | Friends | A boy talks about doing everything with his friends and sticking together with them. Asks "what's your anti-drug?" |
| | Music/Mix Tapes | Animation of youth walking around city streets, listening to music. Youth encounters negative drug influences but continues listening to the music. Youth states that music is his anti-drug. |
| | Swimming | A girl talks about how much she enjoys swimming. Asks "what's your anti-drug?" |
| | What I Need | A youth is confronted by an older teen selling drugs about "what he needs." The youth rattles off a series of positives that he needs in his life. The last positive need is for the dealer to leave him alone. |
| Hispanic | Music/Mix Tapes (Spanish) | Animation of youth walking around city streets, listening to music. Youth encounters negative drug influences but continues listening to the music. Youth states that music is his anti-drug. |
| | Second Trip (Spanish) | Youth are shown skate boarding, climbing, kick boxing, performing in a band. The best kinds of highs come from doing things well, not using drugs. |
| | You Know How to Say It (Spanish) | A youth is offered vegetables, asked to copy homework, asked to ditch basketball, asked to smoke marijuana. "You know how to say no." |

# Wave 3 (continued)

## Table D-1d. Radio Advertisements Played for Youth

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Alberto | Young male talks about why drugs don't go with making music. Music is the anti-drug for this youth. |
| | Excuses | Excuses you can give for not smoking marijuana are provided. |
| | Margot | Female youth has a younger friend with a disability and wants to be her role model. Teaching her about life is more important than taking drugs. Her younger friend is her anti-drug. |
| | Orientation | An orientation to middle school life is presented: pizza, science class, recess, kids who smoke marijuana. Say no to drugs and you won't be treated like a little kid. |
| | The Rant | Ad talks about the lies associated with ecstasy when viewed by nonusers. |
| | What to Say Boy | A friend wants you to smoke "that wacky weed." What do you say? "I get high above the rim." |
| | What to Say Girl | The guy is great, but he wants you to get high. What do you say? "I'd rather go to math camp." |
| | What's Yours | Girl (boy for Black youth) asks "What's your thing? What do you do instead of drugs?" That's your anti-drug. Talks about posting your anti-drug to "whatsyourantidrug.com" or calling 877-979-6300. |
| African American | Alberto | Young male talks about why drugs don't go with making music. Music is the anti-drug for this youth. |
| | What to Say Boy | A friend wants you to smoke "that wacky weed." What do you say? "I get high above the rim." |
| | What to Say Girl | The guy is great, but he wants you to get high. What do you say? "I'd rather go to math camp." |
| | What's Yours | Girl (boy for Black youth) asks "What's your thing? What do you do instead of drugs?" That's your anti-drug. Talks about posting your anti-drug to "whatsyourantidrug.com" or calling 877-979-6300. |
| Hispanic | Jose (Spanish) | Jose is a teen whose anti-drug is music. He sings part of a song called "La Rosa" in the ad. |
| | She Did It (Spanish) | Girls talk to popular girl who says no to marijuana and is still popular. |
| | The First Time (Spanish) | Kids talk about saying no to marijuana for the first time. |

# Wave 2

## Table D-2a. Television Advertisements Shown to Parents

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Clinic | A father and son are shown walking through a clinic – like setting, but finally arrive at a basketball clinic. The ad offers a telephone number to get a book on parent – child activities. |
| | Differences – Drugs | Drugs to 6th grader is medicine; drugs to 7th grader is bag of marijuana. "What a difference a year makes." |
| | Differences – Roach | A roach to a 6th grader is an insect; a roach to 7th grader is part of a marijuana joint. "What a difference a year makes." |
| | Differences – Pipe | A pipe to a 6th grader is plumbing; a pipe to a 7th grader is a marijuana pipe. "What a difference a year makes." |
| | Differences – Weed | A weed to 6th grader is a dandelion; weed to 7th grader is marijuana. "What a difference a year makes." |
| | Instructions – Involved | A girl is shown walking with books, a boy is fixing his bike, a girl is playing with a soccer ball. All have parenting "instructions" visible on their bodies. Wouldn't it be great if kids came with instructions? The instructions advise the parent to stay involved with the child. |
| | Instructions – Reward | Kids are shown playing with their father, eating ice cream, walking. All have parenting "instructions" visible on their bodies. Wouldn't it be great if kids came with instructions? The instructions advise to reward child, provide positive reinforcement. |
| | Instructions – Reward | Kids are shown walking, playing with a dog, running through the hose. All have parenting "instructions" visible on their bodies. Wouldn't it be great if kids came with instructions? The instructions advise to reward child and provide positive reinforcement. |
| African American | Clinic | A father and son are shown walking through a clinic – like setting, but finally arrive at a basketball clinic. The ad offers a telephone number to get a book on parent – child activities. |
| | Instructions – Involved | A boy is shown on a dock, a girl plays with a soccer ball, a boy looks in a mirror. All have parenting "instructions" visible on their bodies. Wouldn't it be great if kids came with instructions? The instructions advise the parent to stay involved with the child. |
| | Instructions – Reward | Kids are shown playing with their father, eating ice cream, walking. All have parenting "instructions" visible on their bodies. Wouldn't it be great if kids came with instructions? The instructions advise to reward child, provide positive reinforcement. |
| | Symptoms | A mother is shown looking depressed, the father is yelling, a young child is curled up in the corner, looking scared. These are the family "symptoms" of teen drug use. |

# Wave 2 (continued)

## Table D-2a. Television Advertisements Shown to Parents (continued)

| Target Audience | Ad name | Description |
|---|---|---|
| Hispanic | Heroes: Dancing (Spanish) | A mother takes her daughter to dance lessons, then watches her daughter's dance recital when the daughter is older. The mother remains the child's hero throughout her life. "Get close to her. . Support her. . .this will help her stay away from drugs." |
| | Heroes: Swimming (Spanish) | A father carries his son as a child, then watches his son's swim meet when he's older. The father remains the child's hero throughout his life. "Get involved in his activities. . . This will help him stay away from drugs." |
| | Mirrors – (Spanish) | A boy wanders through a house of mirrors while his parents search for him. "Your child can be under the illusion that smoking marijuana is harmless." It isn't. |

# Wave 2 (continued)

## Table D-2b. Radio Advertisements Played for Parents

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Desperate | Ad opens with what sounds like a parent lecturing the son about the dangers of drugs. However, the parent is actually playing a video game with the youth and spending time with him. Phone number and web site is given for information about keeping youths off drugs. |
| | Differences – Bag | A bag to a 6th grader is a lunch bag; a bag to a 7th grader is a bag of marijuana. "What a difference a year makes." |
| | Differences – Clip | A clip to a 6th grader is a paper clip; a clip to a 7th grader is a roach clip. "What a difference a year makes." |
| | Keep Trying | A boy describes all the times he was told by his parent to keep trying. He encourages parents to "keep trying" to talk to kids about marijuana. |
| African American | Keep Trying | A boy describes all the times he was told by his parent to keep trying. He encourages parents to "keep trying" to talk to kids about marijuana. |
| | Symptoms | Ad talks about the negative ripple effects that occur in the family when a member is using marijuana. Examples include depression, withdrawal, and hostility. |
| Hispanic | Pepperoni (Spanish) | The best way to keep youth younger than 15 from using drugs is by supervising them and being an effective parent. |

# Wave 2 (continued)

## Table D-2c. Television Advertisements Shown to Youth

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Brothers | A little brother imitates his big brother. The big brother is offered marijuana, but refuses it because he knows he's a role model. |
| | Dance | Animation of a girl dancing to music on her radio. While dancing, she is offered drugs by two boys. She refuses the offer and states that dancing is her anti-drug. |
| | DJ | A boy talks about his feelings when he performs as a disk jockey. Asks "what's your anti-drug?" |
| | Drugs Kill Dreams | Tennis champions Venus and Serena Williams advise against drug use. "Drugs kill dreams." |
| | Hockey | A boy plays hockey without protective gear. Smoking marijuana is like playing hockey without the right equipment. You can't get in the game. |
| | Family | A girl talks about her attachment to her mother. Asks "what's your anti-drug?" |
| | Football | A football player talks about catching a pass. Asks "what's your anti-drug?" |
| | Friends | A boy talks about doing everything with his friends and sticking together with them. Asks "what's your anti-drug?" |
| | How to Say No | Alternative ways (angry, rap, dramatic) to say no to drugs are shown. |
| | Icon | Ad shows a collage of images of various activities. Asks "what's your anti-drug?" |
| | Love | A girl talks about the love she feels for her cat. Asks "what's your anti-drug?" |
| | Mary J. Blige | Singer Mary J. Blige talks about loving and accepting yourself and staying drug free. |
| | Mother/Daughter | A mother talks about how proud she is of her daughter. The daughter meets her friend in the park to smoke marijuana. "Smoking marijuana won't kill you, but it will kill your mother." |
| | No Thanks | A boy at a party is offered marijuana. Different ways to say no to drugs are shown. |
| | Swimming | A girl talks about how much she enjoys swimming. Asks "what's your anti-drug?" |
| | Tara Lipinski | Important female sports figures in past paved the way for women today to play sports. Figure skating champion Tara Lipinski is featured and counsels against drug use. |
| | U.S. Women's Soccer Team | The members of the 1999 World Champion U.S. Women's Soccer Team talk about what a great time it is to be a girl. "Don't blow it by getting involved with drugs." |
| | Vision Warrior | Young man talks about how smoking marijuana led him to use harder drugs. |

# Wave 2 (continued)

## Table D-2c. Television Advertisements Shown to Youth (continued)

| Target Audience | Ad name | Description |
|---|---|---|
| African American | DJ | A boy talks about his feelings when he performs as a disk jockey. Asks "what's your anti-drug?" |
| | Drugs Kill Dreams | Tennis champions Venus and Serena Williams advise against drug use. "Drugs kill dreams." |
| | Family | A girl talks about her attachment to her mother. Asks "what's your anti-drug?" |
| | Football | A football player talks about catching a pass. Asks "what's your anti-drug?" |
| | Friends | A boy talks about doing everything with his friends and sticking together with them. Asks "what's your anti-drug?" |
| | How to Say No | Alternative ways (angry, rap, dramatic) to say no to drugs are shown. |
| | Love | A girl talks about the love she feels for her cat. Asks "what's your anti-drug?" |
| | Mary J. Blige | Singer Mary J. Blige talks about loving and accepting yourself and staying drug free. |
| Hispanic | Most Teens | Girls are shown jumping rope, boxing, playing basketball, and not using drugs. "I'm too smart to be doing stupid stuff like that." |
| | Mother/Daughter | A mother talks about how proud she is of her daughter. The daughter meets her friend in the park to smoke marijuana. "Smoking marijuana won't kill you, but it will kill your mother." |
| | No Skills | Kids are shown making mistakes and unable to play sports well after using drugs. |
| | No Thanks | A boy at a party is offered marijuana. Different ways to say no to drugs are shown. |
| | Swimming | A girl talks about how much she enjoys swimming. Asks "what's your anti-drug?" |
| | Vision Warrior | Young man talks about how smoking marijuana led him to use harder drugs. |
| | Second Trip (Spanish) | Youth are shown skate boarding, climbing, kick boxing, performing in a band. The best kinds of highs come from doing things well, not using drugs. |
| | You Know How to Say It (Spanish) | A youth is offered vegetables, asked to copy homework, asked to ditch basketball, asked to smoke marijuana. "You know how to say no." |

# Wave 2 (continued)

## Table D-2d. Radio Advertisements Played for Youth

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Alberto | Young male talks about why drugs don't go with making music. Music is the anti-drug for this youth. |
| | Excuses | Excuses you can give for not smoking marijuana are provided. |
| | Make You Think | Marijuana makes you think you're interesting and attractive, when you're really not. |
| | Margot | Female youth has a younger friend with a disability and wants to be her role model. Teaching her about life is more important than taking drugs. Her younger friend is her anti-drug. |
| | Orientation | An orientation to middle school life is presented: pizza, science class, recess, kids who smoke marijuana. Say no to drugs and you won't be treated like a little kid. |
| | What to Say Boy | A friend wants you to smoke "that wacky weed." What do you say? "I get high above the rim." |
| | What to Say Girl | The guy is great, but he wants you to get high. What do you say? "I'd rather go to math camp." |
| | What's Yours | Girl (boy for Black youth) asks "What's your thing? What do you do instead of drugs?" That's your anti-drug. Talks about posting your anti-drug to "whatsyourantidrug.com" or calling 877-979-6300. |
| African American | Alberto | Young male talks about why drugs don't go with making music. Music is the anti-drug for this youth. |
| | If Pot Were a Person | Reasons are given why, if pot were a person, you wouldn't like him. He'd make you quit sports, get you in trouble with your parents. |
| | Mary J. Blige | Singer Mary J. Blige talks about loving and accepting yourself and staying drug free. |
| | Money | Items are listed that you can buy with your money if you don't buy marijuana. |
| | What to Say Boy | A friend wants you to smoke "that wacky weed." What do you say? "I get high above the rim." |
| | What to Say Girl | The guy is great, but he wants you to get high. What do you say? "I'd rather go to math camp." |
| | What's Yours | Girl (boy for Black youth) asks "What's your thing? What do you do instead of drugs?" That's your anti-drug. Talks about posting your anti-drug to "whatsyourantidrug.com" or calling 877-979-6300. |

# Wave 2 (continued)

## Table D-2d. Radio Advertisements Played for Youth (continued)

| Target Audience | Ad name | Description |
|---|---|---|
| Hispanic | Boy Meets Girl (Spanish) | A boy who uses drugs meets girl he's interested in. He thinks he's making a good impression, but she thinks he's a loser. |
| | She Did It (Spanish) | Girls talk to popular girl who says no to marijuana and is still popular. |
| | The First Time (Spanish) | Kids talk about saying no to marijuana for the first time. |
| | Typical Story (Spanish) | A boy's friends tell him to try smoking marijuana. He says he doesn't want to smoke. They insist. He says, "I don't need that." |
| | Weekend (Spanish) | A young man laughs and rambles incoherently when friends ask him about his "incredible" weekend. He thinks his story is great. But they can't understand anything he says. |

# Wave 1

## Table D-3a. Television Advertisements Shown to Parents

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Differences – Drugs | Drugs to 6th grader is medicine; drugs to 7th grader is bag of marijuana. "What a difference a year makes." |
| | Differences – Pipe | A pipe to a 6th grader is plumbing; a pipe to a 7th grader is a marijuana pipe. "What a difference a year makes." |
| | Differences – Pot | Pot to a 6th grader is a flower pot; pot to a 7th grader is marijuana. "What a difference a year makes." |
| | Differences – Roach | A roach to a 6th grader is an insect; a roach to 7th grader is part of a marijuana joint. "What a difference a year makes." |
| | Differences – Weed | A weed to 6th grader is a dandelion; weed to 7th grader is marijuana. "What a difference a year makes." |
| | Drugs Kill Dreams | Tennis champions Venus and Serena Williams advise against drug use. "Drugs kill dreams." |
| | Email | A father types an email on his computer while his child plays video game in the background. Spending time with your kids is most effective deterrent to drug use. "Could you send one less email?" |
| | Funeral | Mortuary employees talk about the realities of planning funerals for young people. The ad captions discuss the risk of death from using inhalants. |
| | Office | A typical office is shown at 5:00 PM. Be aware of at-risk times—5:00 PM is the time kids are most likely to be offered drugs. Be sure to check in with them. |
| | Phone | A mother talks on the kitchen phone while child sits in background looking bored. Spending time with your kids is the most effective drug deterrent. "Could you make one less call?" |
| | Symptoms | A mother is shown looking depressed, the father is yelling, a young child is curled up in the corner, looking scared. These are the family "symptoms" of teen drug use. |
| | TV | A father watches TV show while his daughter skims a magazine on the couch. Kids who are younger than 15 and using marijuana are more likely to use other drugs. Spending time with your kids is the most effective deterrent to drug use. "Why do we watch so much television?" |
| | Under Your Nose | Camera pans through house showing everyday items that kids sniff to get high. Parents are unaware of the dangers of sniffing everyday household products. |
| African American | Drugs Kill Dreams | Tennis champions Venus and Serena Williams advise against drug use. "Drugs kill dreams." |
| | Office | A typical office is shown at 5:00 PM. Be aware of at-risk times—5:00 PM is the time kids are most likely to be offered drugs. Be sure to check in with them. |
| | Symptoms | A mother is shown looking depressed, the father is yelling, a young child is curled up in the corner, looking scared. These are the family "symptoms" of teen drug use. |

# Wave 1 (continued)

## Table D-3a. Television Advertisements Shown to Parents (continued)

| Target Audience | Ad name | Description |
|---|---|---|
| Hispanic | Game Show (Spanish) | A parent-child game show is shown. The mother knows where Mozart was born. But her child knows about marijuana. Parents would be surprised about what their kids know about marijuana. |
| | Heroes: Dancing (Spanish) | A mother takes her daughter to dance lessons, then watches her daughter's dance recital when the daughter is older. The mother remains the child's hero throughout her life. "Get close to her. . Support her. . .this will help her stay away from drugs." |
| | Heroes: Swimming (Spanish) | A father carries his son as a child, then watches his son's swim meet when he's older. The father remains the child's hero throughout his life. "Get involved in his activities. . . This will help him stay away from drugs." |
| | Phone (Spanish) | A mother talks on the kitchen phone while child sits in background looking bored. Spending time with your kids is the most effective drug deterrent. "Could you make one less call?" |
| | Under Your Nose (Spanish) | Camera pans through house showing everyday items that kids sniff to get high. Parents are unaware of the dangers of sniffing everyday household products. |

# Wave 1 (continued)

## Table D-3b. Radio Advertisements Played for Parents

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Basketball | Activities are listed that kids would rather do than drugs. The number one deterrent to drugs is parents and the time spent with their kids. |
| | Cooking Dinner | Boredom is one reason kids get involved with drugs. Stay involved with your kids. |
| | Differences – Bag | To a 6th grader, a bag is something that holds your lunch; to a 7th grader, it's something that holds your marijuana. "What a difference a year makes." |
| | Differences – Grass | To a 6th grader, grass is something you cut; to a 7th grader, it's something you smoke. "What a difference a year makes." |
| | Happy Birthday Steven | A mother describes what she does (feeding, bathing) to take care of her teenaged son who used inhalants and suffered brain damage. |
| | Keep Trying | A boy describes all the times he was told by his parent to keep trying. He encourages parents to "keep trying" to talk to kids about marijuana. |
| | Tree Fort | Activities are suggested to do with your kids: rollerblade, play chess, go to movie. Be aware of at-risk hours—between 4 pm and 6 pm is when kids are most likely to try drugs. |
| African American | Keep Trying | A boy describes all the times he was told by his parent to keep trying. He encourages parents to "keep trying" to talk to kids about marijuana. |
| Hispanic | Game Show (Spanish) | A parent-child game show is shown. The mother knows where Mozart was born. But her child knows about marijuana. Parents would be surprised about what their kids know about marijuana. |
| | Happy Birthday Raoul (Spanish) | A mother describes what she does (feeding, bathing) to take care of her teenaged son who used inhalants and suffered brain damage. |
| | Pepperoni (Spanish) | The best way to keep youth younger than 15 from using drugs is by supervising them and being an effective parent. |

# Wave 1 (continued)

## Table D-3c. Television Advertisements Shown to Youth

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Andy McDonald | Skate boarding champion Andy McDonald talks about getting high from skate boarding, not drugs. |
| | Brothers | A little brother imitates his big brother. The big brother is offered marijuana, but refuses it because he knows he's a role model. |
| | Dixie Chicks | The band, the Dixie Chicks, talk about the temptations to use drugs and advise against drug use. |
| | How to Say No | Alternative ways (angry, rap, dramatic) to say no to drugs are shown. |
| | Michael Johnson | Michael Johnson, the world's fastest 200m and 400m runner, is featured. "None of this would be possible if I had used drugs." |
| | No Thanks | A boy at a party is offered marijuana. Different ways to say no to drugs are shown. |
| | Scatman | Scatman performs in a music video style to convey that "Drugs ain't about nothing." |
| African American | Drugs Kill Dreams | Tennis champions Venus and Serena Williams advise against drug use. "Drugs kill dreams." |
| | How to Say No | Alternative ways (angry, rap, dramatic) to say no to drugs are shown. |
| | Most Teens | Girls are shown jumping rope, boxing, playing basketball, and not using drugs. "I'm too smart to be doing stupid stuff like that." |
| | Venus and Serena Williams | Tennis champions Venus and Serena Williams advise against drug use. "Drugs kill dreams." |
| Hispanic | Fast Food (Spanish) | A young boy under the influence of drugs can't answer when asked what he wants at a fast food restaurant. He is ridiculed by others in line and embarrasses himself. |
| | Natural High (Spanish) | Youth are shown skate boarding, climbing, kick boxing, performing in a band. The best kinds of highs come from doing things well, not using drugs. |
| | Second Trip (Spanish) | Youth are shown skate boarding, climbing, kick boxing, performing in a band. The best kinds of highs come from doing things well, not using drugs. |
| | You Know How to Say It (Spanish) | A youth is offered vegetables, asked to copy homework, asked to ditch basketball, asked to smoke marijuana. "You know how to say no." |
| | Test (Spanish) | A young girl under the influence of drugs doodles on a test and can't answer any of the questions. She disappoints the teacher and herself. |

# Wave 1 (continued)

## Table D-3d. Radio Advertisements Played for Youth

| Target Audience | Ad name | Description |
|---|---|---|
| General Market | Brother Jeff | The things that older brother Jeff can do are featured. Jeff doesn't get high because he knows his little brother looks up to him. |
| | Excuses | Excuses you can give for not smoking marijuana are provided. |
| | Make You Think | Marijuana makes you think you're interesting and attractive, when you're really not. |
| | Orientation | An orientation to middle school life is presented: pizza, science class, recess, kids who smoke marijuana. Say no to drugs and you won't be treated like a little kid. |
| | Scatman | Scatman performs in a music video style to convey that "Drugs ain't about nothing." |
| | Stressed | Girls talk about who is stressed out and who has it the worst. But the girl using drugs is really the one who's doing worst. |
| | What to Say Boy | A friend wants you to smoke "that wacky weed." What do you say? "I get high above the rim." |
| | What to Say Girl | The guy is great, but he wants you to get high. What do you say? "I'd rather go to math camp." |
| African American | If Pot Were a Person | Reasons are given why, if pot were a person, you wouldn't like him. He'd make you quit sports, get you in trouble with your parents. |
| | Kathy and Jackie | Kathy talks about her best friend Jackie and how, if they got high, they wouldn't have fun together |
| | Money | Items are listed that you can buy with your money if you don't buy marijuana. |
| | Steven | An urban youth talks about seeing a drug bust on Thanksgiving, being happy, staying true to himself and drug free. |
| | What I Don't Do | A rap song is played that conveys the message that I don't do drugs and it will be all right. |
| | What to Say Boy | A friend wants you to smoke "that wacky weed." What do you say? "I get high above the rim." |
| | What to Say Girl | The guy is great, but he wants you to get high. What do you say? "I'd rather go to math camp." |

# Wave 1 (continued)

## Table D-3d. Radio Advertisements Played for Youth (continued)

| Target Audience | Ad name | Description |
|---|---|---|
| Hispanic | Boy Meets Girl (Spanish) | A boy who uses drugs meets girl he's interested in. He thinks he's making a good impression, but she thinks he's a loser. |
| | Laugh (Spanish) | Boy who is high can't stop laughing long enough to finish the story he's trying to tell. |
| | She Did It (Spanish) | Girls talk to popular girl who says no to marijuana and is still popular. |
| | The First Time (Spanish) | Kids talk about saying no to marijuana for the first time. |
| | Typical Story (Spanish) | A boy's friends tell him to try smoking marijuana. He says he doesn't want to smoke. They insist. He says, "I don't need that." |
| | Weekend (Spanish) | A young man laughs and rambles incoherently when friends ask him about his "incredible" weekend. He thinks his story is great. But they can't understand anything he says. |

# Appendix E
# Construction of Exposure and Outcome Indices

There are two types of indices used in this report, exposure indices and outcome indices. The algorithm for calculating the exposure indices is nearly the same as that described in the second semi-annual report. Sections E.1 and E.2 basically repeat the documentation of these indices. The general exposure index is documented in Section E.1 and the specific in E.2.[1] Section E.3 covers the process for imputation of ad-level recall. This section is more detailed than in the last report. (The procedure itself is little changed.)

The outcome indices are new. They are explained in Section E.4.

## E.1 General Exposure Index

One index is a "general exposure" index (GEI) based on questions D10-D12 of the youth and child questionnaires and on questions F1-F4 of the parent questionnaire. The GEI captures exposure through a very wide variety of channels as can be seen by examining the parent questions in Figure E-1 on page E-2. Note that in each question, the reference period is "in recent months." The questions for youth are completely parallel.

The responses to these questions are combined in a way that is meant to reflect the total number of ad viewings experienced by the respondent. Each possible response was translated into a certain number of viewings over a 1-month period, as shown in Table E-1, assuming that the average person would mostly refer to the last month in trying to interpret "recent months." The four responses were then added together to create a variable running from 0 to a maximum of 180. This continuous scale was split at the values of 4 and 12, as shown in Table E-2. The categories in Table E-2 were chosen to be easy to communicate and also to induce a reasonable distribution of the sample. This was important because too small of a sample in the low exposure group would lead to unacceptably unstable estimates of direct effects.

**Table E-1. Coding of general exposure questions**

| Response Category | New Value |
|---|---|
| Not at all | 0 |
| Less than 1 time a month | 0.5 |
| 1 to 3 times a month | 2 |
| 1 to 3 times a week | 8 |
| Daily or almost daily | 30 |
| More than 1 time a day | 45 |

[1] Section F.3 of the second semi-annual report consists of a rationale for the construction of two indices rather than a single index. That material is not repeated here.

**Table E-2. Cutpoints for GEI**

| Lower bound in GEI | Upper bound in GEI | New value for categorical version | Recode Label |
|---|---|---|---|
| 0 | 3.999 | 1 | Low: Less than 4 times per month |
| 4 | 11.999 | 2 | Medium: 4 to less than 12 times per month |
| 12 | $\infty$ | 3 | High: 12 or more times per month |

**Figure E-1. Parent Questions on General Exposure**

The next questions ask about anti-drug commercials or "ads" that are intended to discourage ***illicit drug*** use.

F1. In recent months, about how often have you seen such anti-drug ads on TV, or heard them on the radio?

Not at all ..... 1
Less than one time a month ..... 2
1 to 3 times a month..... 3
1 to 3 times a week..... 4
Daily or almost daily ..... 5
More than 1 time a day ..... 6

F2. In recent months, about how often have you seen such anti-drug ads in newspapers or magazines?

Not at all ..... 1
Less than one time a month ..... 2
1 to 3 times a month..... 3
1 to 3 times a week..... 4
Daily or almost daily ..... 5
More than 1 time a day ..... 6

F4. In recent months, about how often have you seen any anti-drug billboards or other public anti-drug ads such as on buses, in malls, or at sports events?

Not at all ..... 1
Less than one time a month ..... 2
1 to 3 times a month..... 3
1 to 3 times a week..... 4
Daily or almost daily ..... 5
More than 1 time a day ..... 6

F3. In recent months, about how often have you seen such anti-drug ads in the movie theaters or on rental videos?

Haven't gone to movies or rented videos in recent months ..... 0
Not at all ..... 1
Less than 1 time a month..... 2
1 to 3 times a month..... 3
1 to 3 times a week..... 4
Daily or almost daily ..... 5
More than 1 time a day ..... 6

## E.2 Recall Aided-Exposure Index

The second index is a "recall-aided exposure" index (RAEI) based on the specific TV and radio ads available for sampling. For parents, exposures to TV and radio ads are combined. For youth, only TV exposure is used.* As discussed in Chapters 2 and 3, a selection of ads projected to be on the air in the two calendar months preceding the month of interview were played for respondents. Ads that were eligible for selection but not actually selected for a particular respondent received imputed responses. The imputation procedures are documented in Section E.4.

After imputation, answers were available to the questions shown in Figure E-2 for every ad that had been on the air in the 60 days preceding the day of interview and that were targeted to the respondent. (This means that for parents, only parent ads were sampled/imputed; for youth, only youth ads were sampled/imputed; for English speakers, only English ads were sampled/imputed; and for Spanish speakers, only Spanish ads were sampled/imputed unless they were bilingual, in which case, ads in both languages were sampled and imputed.)

After imputation, the responses were recoded as shown in Table E-3. These recoded values were then summed across ads to get a total number of viewings. For parents, responses to these questions on both TV and radio ads were summed together. For youth, only responses to the TV ads were summed. After summation, the resulting scales were broken into the categories shown in Table E-4. Four levels were chosen for this index instead of the three chosen for the general index because there was a large sample in the bottom group; the direct effects are more compelling when the low exposure group has extremely low exposure.

## E.3 Ad Imputation Procedures

As explained in Section E.2, only a sample of the on-air ads were actually selected for each respondent. In order to characterize each respondent's total exposure to all ads on the air for the RAEI, it was necessary to impute viewing levels of the nonsample ads. Because different ad sampling rules were used for minorities, and because of the variations in the GRPs of the ads, developing a satisfactory analysis procedure was difficult. Simply summing the recall of the sampled ads would have made minorities appear to have been more heavily exposed because they were shown more ads. Simply averaging the recall of the sampled ads would have made people who were shown ads with low GRP appear to be less heavily exposed than those who were shown ads with high GRP. A weighting approach did not appear feasible because we needed to have a single number for each person to conduct this dose-response analysis. Therefore, imputation appeared to be the simplest and, indeed, the only sensible approach. The imputation does tend to reduce the variation in exposure across people—a fact that is not important for the dose-response relationship. The main concern was to get the best possible ordering of people by exposure. Because we controlled on the general recall of TV and radio ads, we believe the imputation produced a better ordering than simple averages would have done. Several different imputation procedures were used depending on the total number of times that an ad was sampled during a wave. The three procedures were single-cell hotdeck imputation,

* See Section 3.1.4 for a discussion of the rationale for this decision at Wave 1. Once the decision had been made at Wave 1, the algorithm for the index was held steady to allow comparisons with Wave 1.

**Figure E-2. Specific ad questions**

F12a. Now we will show some ads that might or might not have been playing on television around here. Have you ever seen or heard this ad? (PLAY TV AD.)

| | | |
|---|---|---|
| Yes | 1 | |
| No | 2 | (F13a) |
| REFUSED | | (F13a) |
| DON'T KNOW | | (F13a) |

F12b. In recent months, how many times have you seen or heard this ad?

| | | |
|---|---|---|
| Not at all | 1 | (F13a) |
| Once | 2 | |
| 2 to 4 times | 3 | |
| 5 to 10 times | 4 | |
| More than 10 times | 5 | |

**Table E-3. Recoding of Responses to Exposure to Specific Ads**

| Question: Here is another TV ad. Have you ever seen or heard this ad? | [If yes,] In recent months, how many times have you seen or heard this ad? | Recoded Response |
|---|---|---|
| No | | 0.0 |
| Don't know | | 0.5 |
| Yes | Not at all | 0.0 |
| Yes | Once | 1.0 |
| Yes | 2 to 4 times | 3.0 |
| Yes | 5 to 10 times | 7.5 |
| Yes | More than 10 times | 12.5 |

**Table E-4. Cutpoints for RAEI**

| Lower bound in RAEI | Upper bound in RAEI | New value for categorical version | Recode Label |
|---|---|---|---|
| 0 | 1.999 | 0 | None |
| 2 | 7.999 | 1 | One to less than 4 times per month (low) |
| 8 | 23.999 | 2 | 4 to less than 12 times per month (medium) |
| 24 | $\infty$ (90 actual upper limit) | 3 | 12 or more times per month (high) |

n-cell hotdeck imputation, and a MART-based procedure, each of which is explained below. For Wave 3, the single-cell hotdeck was used for 13 of the ads, the n-cell for 15, and the MART-based approach on 13.

## E.3.1 Single-Cell Hotdeck Imputation

This procedure was used whenever the total number of subjects for which an ad was in-scope during a wave was 150 or less. In this situation where there was little information available about the distribution of viewing in the population, the judgment was made that it was best to select a random respondent among those for whom the ad was sampled and then to transcribe the results from the "donor" to the "beggar." The only restrictions on donor choice were that (1) both interviews had to be conducted at times such that the ad in question had been on the air within the 60 days preceding the

interview and (2) both donor and beggar consume the medium in the language of the ad (English or Spanish).

## E.3.2 N-Cell Hotdeck Imputation

When there was more information about the distribution of viewing of an ad (sample size between 151 and 499), more complex procedures were used to match donors and beggars. In addition to matching on eligibility for the ad (on air in preceding 60 days and right language), matching was done on the length of time the ad had been on the air (3 categories), whether the respondent's home had cable/satellite service, and the level of general recall of drug-related advertisements on TV and radio. If perfect matching on all three criteria was impossible, the software had an automatic feature that searched for a suitable donor by relaxing the match criteria. The criteria are relaxed according to a predetermined order fixed by the user. In this case, general recall was relaxed first when necessary.

## E.3.3 MART-Based Imputation

MART (Multiple Additive Regression Trees) is an iterative method that may be used to form predictive models for variables of interest. It is particularly well suited for modeling ordinal and multilevel categorical variables in terms of other ordinal and multilevel categorical variables. It was designed to handle large numbers of potential predictor variables in a largely automated fashion, requiring little human intervention. Guidance from the developer of the software (Jerome Friedman of Stanford) indicated that at least 500 observations were required for satisfactory performance. That is why the hotdeck procedures were used for the less frequently sampled ads. When there is adequate sample size, the advantage of using MART is that the procedure will preserve associations of exposure with a larger set of covariates than is possible with the hotdeck procedures. With the hotdeck, it is only possible to preserve associations with a few characteristics selected prior to the end of data collection. From a large set of potential predictor variables, MART builds a model for the variable of interest (ad viewing in this application) that fits it very closely. In heuristic terms, the difference is somewhat like buying a ready-made suit by mail order versus having one tailor made—except in this case, the tailor is a robot.

MART was used to form models for viewing the various ads. These models were then used to impute ad viewing. MART delivers the model in the form of a probability distribution on the categories for each respondent. In other words, MART calculates, for each respondent and for every possible response value, the chance of the actual response being the given value. This probability distribution may then be used to impute an unknown response.

The MART procedure calculates the response distribution for every respondent based on data on known responses and possible predictor variables. It relates responses to predictors using the respondents for whom responses are available. It then uses these relationships to calculate the likelihood of different responses for the respondents for whom the response is missing. It is an extremely flexible method, which can accommodate large numbers of possible predictors and possibly complicated interactions between them. It can also deal with missing predictor values by treating predictor nonresponse as a different category of the predictor.

In our use of MART for imputing responses to the different advertisements shown, we have used data that has been pooled over advertisements. This allows MART to exploit the similarity of the predictor-response relationship across advertisements and thus "borrow strength" during calculation.

However, possible individual advertisement characteristics are preserved by adding an advertisement indicator variable as a predictor. This allows MART to distinguish between the different advertisements and prevents individual signal from being overwhelmed by overall signals.

Once the probability distribution is calculated by MART, an impute value is chosen as follows. A random number is generated from a uniform distribution. The MART-predicted probability distribution is cumulated over categories to give break points and the category in which the random number lies is noted. This category is then determined to be the imputed response value. If there are *n* response categories, all of which were determined to be equally likely, this procedure is equivalent to rolling an *n*-faced die to determine the unknown response value. Each face of the die is associated with a particular value of the response. To carry the analogy over in case of the categories being unequally likely, as is usually the case in reality, we would have to imagine the die as being unequally weighted so that the probability of each face corresponds the probability of the response category associated with the face.

Details of the MART procedure may be found in Friedman et al. (2000). A separate paper on the quality of imputations based on MART versus the more traditional hotdeck is under preparation. Testing done at Westat indicated a slight MART advantage in the marginal distributions of the variables being imputed. The more important MART advantage is that associations are preserved with a larger set of covariates. The set of covariates that were fed to MART and may thus be considered to have had their associations with exposure preserved include the following data.

## Parent Data

- Parental response group used to determine which advertisements they were shown during the interview (African American, monolingual English, monolingual Spanish, bilingual English/Spanish)
- Parental TV consumption on weekdays
- Parental TV consumption on weekends
- Primary language of TV viewing by parent
- Availability of cable/satellite TV in the household
- Parental consumption of TV channels focused on African Americans in last 30 days
- Parental consumption of TV channels focused on Hispanics/Latinos in last 30 days
- Parental recall of watching or hearing anti-drug advertisements on TV or radio
- Parental age
- Parental marital status
- Parental educational attainment
- Parent's attendance of religious services
- Importance of religion to the parent

- Parental income
- Parental weekday radio consumption
- Parental weekend radio consumption
- Primary language in which parent listens to radio
- Urbanity of the neighborhood
- Parental race
- Parental gender
- Region of the country
- Number of days out of the last 30 that the advertisement was aired
- Number of days out of the last 60 that the advertisement was aired
- Number of days out of the last 90 that the advertisement was aired
- Indicator of which advertisement is being shown

## Youth Data

- The youth category (teen or child)
- Youth response group used to determine which advertisements they were shown during the interview (African American, monolingual English, monolingual Spanish, bilingual English/Spanish)
- Youth report on school enrollment in previous 12 months
- Youth school grade level
- Average grade in school
- Whether the youth's school was in session in the last 30 days
- Youth report of school days missed due to illness in the last 30 days
- Youth report of cut school days in the last 30 days
- Youth's attendance of religious services
- Importance of religion to the youth
- Youth TV consumption on weekdays
- Youth TV consumption on weekends
- Primary language in which the youth watches TV

- Youth consumption of radio on weekdays
- Youth consumption of radio on weekends
- Primary language in which the youth listens to radio
- Youth consumption of TV channels focused on music in last 30 days
- Youth consumption of TV channels focused on sports in last 30 days
- Youth consumption of TV channels focused on African Americans in last 30 days
- Youth consumption of TV channels focused on Hispanics/Latinos in last 30 days
- Youth recall of watching or hearing anti-drug advertisements on TV or radio
- Youth age
- Availability of cable/satellite TV in the household
- Urbanity of the neighborhood
- Youth race
- Youth gender
- Region of the country
- Number of days out of the last 30 that the advertisement was aired
- Number of days out of the last 60 that the advertisement was aired
- Number of days out of the last 90 that the advertisement was aired
- Indicator of which advertisement is being shown

## E.3.4 Some Evaluative Information on the N-cell Hotdeck Application

As mentioned above, the sample sizes for some ads were too small to make use of MART. Parametric modeling procedures would also have failed on these small sample sizes, in particular given the nonnormality of the recall data. This nonnormality is demonstrated in Table E-5. Kolmogorov-Smirnoff tests were carried out to check how significantly the response distribution differed from he normal distribution. Skew and kurtosis were also calculated and are shown in the table. Clearly, these data are far from normal, so any parametric-based imputation of the ad-level data would be difficult.

Despite this nonnormality, however, it is interesting to use linear modeling as a means to partially demonstrate the process features of the hotdeck. The variables used to match beggars with donors in the n-cell hotdeck were chosen prior to processing of the Wave 1 data. As discussed in Section E.3.2, there were three of these matching variables. Linear models were fit for the ad-level recall data in terms of the three matching variables as a means of confirming that these *a prior* choices for matching variables were reasonable. A separate linear model was fit for each audience and medium (i.e., for

each of parent TV, parent radio, youth TV, and youth radio). Interactions were examined. The results are shown in Table E-6.

**Table E-5. Non-normality of Ad-level Recall Data**

| | Kolmogorov-Smirnoff Test | | Moments of Ad-level Recall Data | | | |
|---|---|---|---|---|---|---|
| Audience and Medium | Statistic | p value | Mean | Standard Deviation | Skewness | Kurtosis |
| Parent TV | 0.3382 | 0.0000 | 2.0026 | 3.5163 | 1.9272 | 5.7102 |
| Parent Radio | 0.4005 | 0.0000 | 1.1680 | 2.6081 | 2.7929 | 10.7849 |
| Youth TV | 0.3194 | 0.0000 | 2.2292 | 3.8177 | 1.7734 | 4.8855 |
| Youth Radio | 0.4233 | 0.0000 | 0.8674 | 2.3569 | 3.5381 | 15.8444 |

Note : A Normal distribution has a skewness of 0 and kurtosis of 3.

**Table E-6. Results of ANOVA Analysis for WESDECK Imputation Procedure**

| | Parent TV Model | | Parent Radio Model | | Youth TV Model | | Youth Radio Model | |
|---|---|---|---|---|---|---|---|---|
| Effect *(Degrees of Freedom)* | F-Statistic | p value | F-Statistic | p value | F-Statistic | p value | F-Statistic | p value |
| Availability of cable TV in the household (TCABLETV) *(1)* | 0.0495 | 0.8239 | 0.0004 | 0.9837 | 4.4984 | **0.0343** | 1.5482 | 0.2136 |
| Level of general recall of drug-related advertising on TV and radio (TVRAD) *(5)* | 24.5390 | **0.0000** | 12.1425 | **0.0000** | 6.9137 | **0.0000** | 7.1031 | **0.0000** |
| Length of time advertisement had been on air in the 60 days preceding the interview - 3 levels (AIR60) *(2)* | 7.1532 | **0.0008** | 3.9412 | **0.0197** | 13.9294 | **0.0002** | 8.0582 | **0.0003** |
| TCABLETV*TVRAD *(5)* | 0.4582 | 0.8075 | 0.6667 | 0.6488 | 1.9909 | *0.0782* | 0.6579 | 0.6555 |
| TCABLETV*AIR60 *(2)* | 2.3608 | *0.0948* | 2.4039 | *0.0908* | 1.1065 | 0.2933 | 0.2748 | 0.7597 |
| TVRAD*AIR60 *(9)* | 0.6350 | 0.7847 | 0.8738 | 0.5482 | 1.3894 | 0.2263 | 0.6370 | 0.7830 |
| TCABLETV*TVRAD*AIR60 *(6)* | 2.2240 | **0.0235** | 2.0710 | *0.0539* | 2.0056 | *0.0922* | 1.0962 | 0.3613 |

Note : **Boldface** denotes effect is significant at 5 percent level. *Underlined Italics* denote effects significant at 10 percent level. Note, however, that since the response variable is highly nonnormal as demonstrated above the significance levels of the ANOVA are highly approximate.

The availability of cable or satellite TV service was not as important as initially guessed it would be, but is still relevant for youth TV. Within each audience and medium, the general level of recall of anti-drug advertisements on TV and radio was highly relevant to recall of specific Campaign-sponsored advertisements. It would, of course, have been surprising not to find this relationship. Similarly, the number of recent weeks during which the ad had been played was extremely important. In several cases, some of the interaction terms were also found to be significant.

## E.4 Outcome Indices

In order to ameliorate problems caused by multiple comparisons, new outcome indices were created for Wave 3 and retrospectively applied to Waves 1 and 2. By focusing on a smaller number of outcomes, the expected number of false positive findings is reduced. In addition, if the outcome indices are well-constructed, it is possible that the index will be more sensitive to change or effects than any of the components individually.

For youth, a total of just four outcome indices were produced. For parents, there were two. These indices are different from scales. Scales are functions of several variables that are thought to measure the same latent construct. Indices are more general functions of several variables, designed with a particular objective in mind. Well-known indices in other fields include the gross domestic product (GDP), the Consumer Price Index (CPI), and various quality of life indices comparing cities.

In this case, the indices were created with the specific objective of predicting a primary cognitive or behavioral outcome. For youth, the primary outcome was the intention not to use a drug in the future. For parents, the primary outcome was either talking with their kids about drugs or monitoring their kids closely. More detail is given below on each set of indices.

## E.4.1 Youth

For youth, the two primary outcomes were intentions to avoid marijuana use and intentions to avoid inhalant use. Referring back to Figure 2-C, intentions are theorized to be influenced by (1) knowledge, beliefs, and attitudes; (2) perceived social norms, and (3) self-efficacy to avoid drug usage. Questionnaire items that corresponded to each of the influential cognition families were used to form parametric models of the primary outcomes. The concept behind this practice was to let the data inform the Evaluation team about which items within a family really were influential on the primary outcome.

For example, in Table E-11, it can be seen that among the self-efficacy items included in the questionnaire, the most important in terms of influencing intentions to avoid marijuana use are feelings of self-efficacy to refuse marijuana when home alone and sad or bored; when on school property, and when hanging out at a friend's house without parents. Kids who are completely sure that they could refuse marijuana when home alone and sad/bored, or when hanging out at a friends house, were much more likely to have strong intentions to avoid future marijuana use. Conversely, youth who were completely sure that they could refuse offers when on school grounds were less likely to have such strong intentions. Feelings of self-efficacy at parties and at the suggestion of close friends do not appear to be influential on intentions for future use.

The indices for beliefs/attitudes and for social norms were more difficult to construct. For these areas, there were skip patterns in the questionnaires that forced part of the sample to answer questions about trial use and forced the balance to answer questions about regular use. The skip patterns were partly random and partly a function of past marijuana use. As a way to use different questions to create a single index that was meaningfully defined on the entire sample, a complex procedure was used to create each index.

The first step in the process was to model intentions to avoid future use on nonusers in terms of beliefs and attitudes about trial use. This model is shown in Table E-7. The second step was to model intentions to avoid future use on nonusers in terms of beliefs and attitudes about regular use. This model is shown in Table E-8. The third step was to shift and rescale these subindices to that they had a common mean and standard deviation on the population of nonusers. The transformed functions were then applied to the questions about regular use asked of users. (Users were never asked about future trial use.) The end result of this operation was to create an index on the entire dataset that reflects the influence on intentions for avoidance of future use of an amalgam of beliefs and attitudes about both marijuana trial and regular marijuana use.

**Table E-7. Model for intentions to avoid any marijuana use among 12 to 18 year old non-marijuana users in terms of personal beliefs and attitudes about trial marijuana use**

| Quex Item | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| C3a(a) | Trying marijuana would upset parents/caregivers | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | 0.1524 | 0.2695 |
| | | 4 | Likely | -0.5901 | 0.3027 |
| | | 5 | Very likely | 0.4377 | 0.2118 |
| C3a(b) | Trying marijuana would cause legal trouble for youth | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | -0.3179 | 0.1949 |
| | | 4 | Likely | 0.1289 | 0.2095 |
| | | 5 | Very likely | 0.1891 | 0.2329 |
| C3a(c) | Trying marijuana would cause youth to lose control | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | -0.1752 | 0.2224 |
| | | 4 | Likely | -0.2441 | 0.2164 |
| | | 5 | Very likely | 0.4193 | 0.3087 |
| C3a(d) | Trying marijuana would cause youth to use stronger drugs | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | -0.0221 | 0.2478 |
| | | 4 | Likely | 0.3056 | 0.2823 |
| | | 5 | Very likely | -0.2835 | 0.3883 |
| C3a(e) | Trying marijuana would cause youth to be more relaxed | 1 | Very unlikely | 0.1361 | 0.2427 |
| | | 2 | Unlikely | 0.0211 | 0.2468 |
| | | 3-5 | Neither likely nor unlikely, likely, or very likely | -0.1572 | 0.2036 |
| C3a(f) | Trying marijuana would cause youth to have a good time with friends | 1 | Very unlikely | 0.4546 | 0.2688 |
| | | 2 | Unlikely | -0.4197 | 0.2310 |
| | | 3-5 | Neither likely nor unlikely, likely or very likely | -0.0349 | 0.2180 |
| C3a(g) | Trying marijuana would cause youth to feel better | 1 | Very unlikely | -0.1994 | 0.2331 |
| | | 2 | Unlikely | 0.1629 | 0.2189 |
| | | 3-5 | Neither likely nor unlikely, likely, or very likely | 0.0365 | 0.2327 |
| C3a(h) | Trying marijuana would cause youth to be like the coolest kids | 1 | Very unlikely | 0.3274 | 0.1942 |
| | | 2 | Unlikely | 0.2613 | 0.2122 |
| | | 3-5 | Neither likely nor unlikely, likely, or very likely | -0.5886 | 0.2038 |
| C4a | Youth perception of trying marijuana in the next year (7-point scale from “extremely bad” to “extremely good”) | 1 | | 1.4258 | 0.2460 |
| | | 2 | | -0.3259 | 0.2440 |
| | | 3 | | -0.2839 | 0.3129 |
| | | 4-7 | | -0.8160 | 0.2806 |
| C5a | Youth perception of trying marijuana in the next year (7-point scale from “extremely unenjoyable” to “extremely enjoyable”) | 1 | | 0.8747 | 0.2433 |
| | | 2 | | 0.2961 | 0.2593 |
| | | 3 | | -0.6307 | 0.2843 |
| | | 4-7 | | -0.5402 | 0.2846 |

**Table E-8. Model for intentions to avoid any marijuana use among 12-18 year old non-marijuana users in terms of personal beliefs and attitudes about regular marijuana use**

| Quex Item | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| C3b(a) | Regular marijuana use would damage youth's brain | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | -0.1549 | 0.2164 |
| | | 4 | Likely | -0.0435 | 0.1858 |
| | | 5 | Very likely | 0.1984 | 0.2141 |
| C3b(b) | Regular marijuana use would mess up youth's life | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | 0.2318 | 0.2415 |
| | | 4 | Likely | -0.0884 | 0.1969 |
| | | 5 | Very likely | -0.1434 | 0.2395 |
| C3b(c) | Regular marijuana use would make youth do worse in school | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | -0.3141 | 0.2464 |
| | | 4 | Likely | -0.0044 | 0.1933 |
| | | 5 | Very likely | 0.3186 | 0.2318 |
| C3b(d) | Regular marijuana use would be acting against youth's moral beliefs | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | -0.2912 | 0.1988 |
| | | 4 | Likely | 0.1467 | 0.1973 |
| | | 5 | Very likely | 0.1446 | 0.2104 |
| C3b(e) | Regular marijuana use would cause youth to lose ambition | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | -0.0250 | 0.2259 |
| | | 4 | Likely | 0.1443 | 0.1977 |
| | | 5 | Very likely | -0.1193 | 0.2447 |
| C3b(f) | Regular marijuana use would cause youth to lose friends' respect | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | -0.5111 | 0.1967 |
| | | 4 | Likely | 0.1517 | 0.1983 |
| | | 5 | Very likely | 0.3594 | 0.2349 |
| C3b(g) | Regular marijuana use would cause youth to have a good time with friends | 1 | Very unlikely | 1.0099 | 0.2677 |
| | | 2 | Unlikely | -0.6336 | 0.2172 |
| | | 3-5 | Neither likely nor unlikely, likely or very likely | -0.3762 | 0.1953 |
| C3b(h) | Regular marijuana use would cause youth to be more creative and imaginative | 1-3 | Very unlikely, unlikely, or neither likely nor unlikely | -0.1549 | 0.2437 |
| | | 4 | Likely | 0.1546 | 0.3294 |
| | | 5 | Very likely | 0.0004 | 0.3749 |
| C4b | Youth perception of regular marijuana use in the next year (7-point scale from "extremely bad" to "extremely good") | 1 | | 0.9698 | 0.2370 |
| | | 2 | | -0.2337 | 0.2386 |
| | | 3 | | -0.7086 | 0.2921 |
| | | 4-7 | | -0.0275 | 0.3042 |
| C5b | Youth perception of regular marijuana use in the next year (7-point scale from "extremely unenjoyable" to "extremely enjoyable") | 1 | | 0.7496 | 0.2271 |
| | | 2 | | -0.1493 | 0.2414 |
| | | 3 | | -0.2438 | 0.2936 |
| | | 4-7 | | -0.3565 | 0.2451 |

A parallel process was used for social norms. Table E-9 has the parameter estimates for the subindex for social norms about trial use. Table E-10 provides the parallel estimates for the subindex for social norms about regular use. Table E-11 provides the model for intentions to avoid any marijuana use among 12- to 18-year-olds in terms of self-efficacy to refuse offers of marijuana.

One index was created for youth to summarize personal beliefs about inhalants. (There were no questionnaire items on attitudes, social norms or self-efficacy with respect to inhalants.) As with marijuana, the importance of each component in the index was determined from the parametric model for intentions to avoid inhalant use in terms of the components. The fitted model is shown in Table E-12. Perceptions of trial risk are related to intentions to avoid future use. Approval of others' trial of inhalants is also related to intentions to avoid future use.

**Table E-9. Model for intentions to avoid any marijuana use among 12 to 18 year old non-marijuana users in terms of perceived social norms about trial marijuana use**

| Quex Item | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| C6a | Youth perception of most important people's reaction to youth trying marijuana | 1 | Strongly disapprove | 0.3815 | 0.2229 |
| | | 2 | Disapprove | -0.4784 | 0.2455 |
| | | 3-5 | Neither approve nor disapprove, approve or strongly approve | 0.0970 | 0.3381 |
| C7a | Youth perception of close friends' reaction to youth trying marijuana | 1 | Strongly disapprove | 1.0315 | 0.1786 |
| | | 2 | Disapprove | -0.0991 | 0.1618 |
| | | 3-5 | Neither approve nor disapprove, approve or strongly approve | -0.9324 | 0.1681 |
| C8a | Youth perception of parents' reaction to youth trying marijuana | 1 | Strongly disapprove | 0.5658 | 0.2729 |
| | | 2 | Disapprove | 0.0545 | 0.3315 |
| | | 3-5 | Neither approve nor disapprove, approve or strongly approve | -0.6203 | 0.4227 |
| C10a | Youth perception of how many friends have tried marijuana | 1-2 | None or a few | 0.3854 | 0.1918 |
| | | 3 | Some | -0.1872 | 0.2012 |
| | | 4-5 | Most or all | -0.1982 | 0.2568 |
| C11 | Youth perception of how many kids in same grade or same age have tried marijuana | 1-2 | None or a few | 0.3894 | 0.1764 |
| | | 3 | Some | -0.1868 | 0.1607 |
| | | 4-5 | Most or all | -0.2026 | 0.2039 |

**Table E-10. Model for intentions to avoid any marijuana use among 12 to 18 year old non-marijuana users in terms of perceived social norms about regular marijuana use**

| Quex Item | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| C6b | Youth perception of most important people's reaction to youth using marijuana regularly | 1 | Strongly disapprove | 0.6495 | 0.2230 |
| | | 2 | Disapprove | -0.2729 | 0.2472 |
| | | 3-5 | Neither approve nor disapprove, approve or strongly approve | -0.3765 | 0.3476 |
| C7b | Youth perception of close friends' reaction to youth using marijuana regularly | 1 | Strongly disapprove | 0.9112 | 0.1844 |
| | | 2 | Disapprove | -0.0951 | 0.1722 |
| | | 3-5 | Neither approve nor disapprove, approve or strongly approve | -0.8160 | 0.1825 |
| C8b | Youth perception of parents' reaction to youth using marijuana regularly | 1 | Strongly disapprove | -0.0445 | 0.2371 |
| | | 2-5 | Disapprove, neither approve or disapprove, approve or strongly approve | 0.0445 | 0.2371 |
| C10b | Youth perception of how many friends have used marijuana regularly | 1-2 | None or a few | 0.2339 | 0.2050 |
| | | 3 | Some | 0.0106 | 0.2192 |
| | | 4-5 | Most or all | -0.2445 | 0.2814 |
| C12 | Youth perception of how many kids in same grade or same age have used marijuana regularly | 1-2 | None or a few | 0.3827 | 0.1874 |
| | | 3 | Some | -0.1066 | 0.1726 |
| | | 4-5 | Most or all | -0.2761 | 0.2353 |

**Table E-11 Model for intentions to avoid any marijuana use among 12- to 18-year-olds in terms of self-efficacy to refuse offers of marijuana.**

| Quex Item | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| C9(a) | Certainty of refusing marijuana when at a party where most people are using it | 1-3 | Somewhat sure, slightly sure, or not at all sure | -0.1805 | 0.1421 |
| | | 4 | Mostly sure | 0.2339 | 0.1130 |
| | | 5 | Completely sure | -0.0535 | 0.1166 |
| C9(b) | Certainty of refusing marijuana when a very close friend suggests using it | 1-3 | Somewhat sure, slightly sure, or not at all sure | -0.0627 | 0.1530 |
| | | 4 | Mostly sure | -0.1604 | 0.1110 |
| | | 5 | Completely sure | 0.2231 | 0.1197 |
| C9(c) | Certainty of refusing marijuana when home alone and feeling sad or bored | 1-3 | Somewhat sure, slightly sure, or not at all sure | -0.6240 | 0.1402 |
| | | 4 | Mostly sure | -0.0458 | 0.1221 |
| | | 5 | Completely sure | 0.6699 | 0.1051 |
| C9(d) | Certainty of refusing marijuana when on school property | 1-3 | Somewhat sure, slightly sure, or not at all sure | 0.6551 | 0.1892 |
| | | 4 | Mostly sure | -0.3183 | 0.1556 |
| | | 5 | Completely sure | -0.3367 | 0.1356 |
| C9(e) | Certainty of refusing marijuana when hanging out at a friend's house whose parents aren't home | 1-3 | Somewhat sure, slightly sure, or not at all sure | -0.8485 | 0.1527 |
| | | 4 | Mostly sure | -0.1478 | 0.1118 |
| | | 5 | Completely sure | 0.9963 | 0.1221 |

**Table E-12 Model for intentions to avoid any inhalant use among 12- to 18-year-olds in terms of personal anti-inhalant beliefs**

| Quex Item | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| C33a(c) | Youth perception of risk of harm when trying inhalants | 1-2 | No or slight risk | -0.3292 | 0.1177 |
| | | 3 | Moderate risk | 0.0600 | 0.1066 |
| | | 4 | Great risk | 0.2692 | 0.1249 |
| C33a(d) | Youth perception of risk of harm when using inhalants regularly | 1-2 | No or slight risk | 0.2185 | 0.1823 |
| | | 3 | Moderate risk | -0.3062 | 0.1339 |
| | | 4 | Great risk | 0.0876 | 0.1328 |
| C33(c) | Youth approval of others trying inhalants | 1 | Strongly disapprove | 1.3941 | 0.1511 |
| | | 2 | Disapprove | -0.1367 | 0.1153 |
| | | 3-5 | Neither approve nor disapprove, approve, or strongly approve | -1.2574 | 0.1330 |
| C33(d) | Youth approval of others using inhalants regularly | 1 | Strongly disapprove | 0.2942 | 0.1249 |
| | | 2 | Disapprove | -0.1642 | 0.1162 |
| | | 3-5 | Neither approve nor disapprove, approve, or strongly approve | -0.1301 | 0.1412 |

## E.4.2 Parents

Two indices were constructed for parents. One summarized information about cognitive variables surrounding the discussion of drugs with their children. The other summarized information about cognitive variables surrounding monitoring of their children. As for youth, models were constructed for primary outcomes in terms of these cognitive variables in order to summarize only the relevant information. Ordinal logistic regressions were used for the modeling.

For discussions about drugs, the primary outcome variable was a scale based on three types of talking behavior. The scale gives a point for each type: (1) two or more general discussions about drugs, (2) at least conversation on the specific topic of family rules or expectations about drug use, and (3) at least conversation on the specific topic of how to avoid drug use. The scale thus runs from 0 to3, with 0 reflecting no discussion and 3 reflecting a pattern of discussions consistent with Campaign objectives. The cognitive variables to be summarized are shown in Table E-13, along with their coefficients.

For monitoring their children, the primary outcome variable was a scale based on three types of monitoring behavior. The scale gives a point for each type: (1) always or almost always knowing what their child is doing when he/she is away from home, (2) always or almost always having a pretty good idea about their child's plans for the coming day, and (3) never allowing their child to spend his/her free time in the afternoons hanging out with friends without adult supervision. The scale thus runs from 0 to3, with 0 reflecting very weak monitoring and 3 reflecting a pattern of monitoring consistent with Campaign objectives. The cognitive variables to be summarized are shown in Table E-14, along with their coefficients.

**Table E-13. Model for Parental talking scale in terms of cognitive variables surrounding discussion of drugs with their children**

| Quex Item | Description of Variable | Values | Value label | Co-efficient | Standard Error |
|---|---|---|---|---|---|
| D2a | Discussing drug use in the next 6 months with my child would be (7-point scale form "extremely bad" to "Extremely good") | 1-4 | | -0.3066 | 0.0976 |
| | | 5 | | -0.1794 | 0.0757 |
| | | 6 | | 0.0913 | 0.0629 |
| | | 7 | | 0.3947 | 0.0617 |
| D2b | Discussing drug use in the next 6 months with my child would be (7-point scale form "extremely unpleasant " to "Extremely pleasant") | 1-4 | | -0.2097 | 0.0581 |
| | | 5 | | -0.0588 | 0.0519 |
| | | 6 | | -0.0395 | 0.0479 |
| | | 7 | | 0.308 | 0.051 |
| D2c | Discussing drug use in the next 6 months with my child would be (7-point scale form "extremely unimportant" to "Extremely important") | 1-4 | | -0.516 | 0.1043 |
| | | 5 | | -0.279 | 0.0823 |
| | | 6 | | 0.2465 | 0.0669 |
| | | 7 | | 0.5484 | 0.0622 |
| D3a | If my child asked me questions about drug use in general, how sure am I that would be able to talk about illicit drug use with that child? | 1-3 | Very unsure, unsure, or neither sure nor unsure | -0.1814 | 0.1046 |
| | | 4 | Sure | 0.0868 | 0.0668 |
| | | 5 | Very Sure | 0.0945 | 0.0659 |
| D3b | If my child asked me questions about me what specific things he/she could do to stay away from drugs, how sure am I that would be able to talk about illicit drug use with that child? | 1-3 | Very unsure, unsure, or neither sure nor unsure | -0.3382 | 0.1076 |
| | | 4 | Sure | 0.0342 | 0.0662 |
| | | 5 | Very Sure | 0.304 | 0.0671 |
| D3c | If my child and I had been having conflicts over other things not related to drugs, and our relationship were tense, how sure am I that would be able to talk about illicit drug use with that child? | 1-3 | Very unsure, unsure, or neither sure nor unsure | -0.1407 | 0.0482 |
| | | 4 | Sure | 0.0714 | 0.039 |
| | | 5 | Very Sure | 0.0693 | 0.0436 |
| D3d | If my child asked me questions about me about my own past use of drugs, how sure am I that would be able to talk about illicit drug use with that child? | 1-3 | Very unsure, unsure, or neither sure nor unsure | -0.0591 | 0.0562 |
| | | 4 | Sure | 0.0146 | 0.0473 |
| | | 5 | Very Sure | 0.0445 | 0.0423 |

**Table E-14. Model for parental monitoring index in terms of personal beliefs regarding monitoring kids' behavior and activities**

| Quex Item | Description of Variable | Values | Value Label | Coefficient | Standard Error |
|---|---|---|---|---|---|
| C6a | Closely monitoring my child's daily activities would be (7-point scale from "extremely bad" to "extremely good") | 1-4 | | -0.8304 | 0.1135 |
| | | 5 | | -0.1358 | 0.0793 |
| | | 6 | | 0.1675 | 0.0705 |
| | | 7 | | 0.7987 | 0.0727 |
| C6b | Closely monitoring my child's daily activities would be (7-point scale from "extremely unpleasant" to "extremely pleasant") | 1-4 | | -0.3743 | 0.0888 |
| | | 5 | | -0.0235 | 0.0656 |
| | | 6 | | 0.1349 | 0.0605 |
| | | 7 | | 0.2628 | 0.0646 |
| C6c | Closely monitoring my child's daily activities would be (7-point scale from "extremely unimportant" to "extremely important") | 1-4 | | 0.0616 | 0.1536 |
| | | 5 | | -0.0482 | 0.1017 |
| | | 6 | | -0.1347 | 0.0857 |
| | | 7 | | 0.1213 | 0.0802 |
| C7a | Closely monitoring my child's daily activities will make it more likely that my child will do well in school | 1-3 | Strongly disagree, disagree, or neither agree nor disagree | -0.0819 | 0.0812 |
| | | 4 | Agree | -0.1007 | 0.0565 |
| | | 5 | Strongly agree | 0.1827 | 0.0617 |
| C7b | Closely monitoring my child's daily activities will make me feel like I am doing my job as a parent | 1-3 | Strongly disagree, disagree, or neither agree nor disagree | 0.1989 | 0.0931 |
| | | 4 | Agree | -0.1064 | 0.0589 |
| | | 5 | Strongly agree | -0.0925 | 0.0664 |
| C7d | Closely monitoring my child's daily activities will make it less likely that my child will try any drug, even once or twice | 1-3 | Strongly disagree, disagree, or neither agree nor disagree | -0.1213 | 0.0712 |
| | | 4 | Agree | -0.1000 | 0.0552 |
| | | 5 | Strongly agree | 0.2212 | 0.0651 |
| C7e | Closely monitoring my child's daily activities will make it less likely that my child will use any drug nearly every month | 1-3 | Strongly disagree, disagree, or neither agree nor disagree | -0.0375 | 0.0725 |
| | | 4 | Agree | -0.0870 | 0.0568 |
| | | 5 | Strongly agree | 0.1245 | 0.0645 |
| C7f | Closely monitoring my child's daily activities will make my child feel I am invading their privacy | 1 | Strongly disagree | 0.3013 | 0.0614 |
| | | 2 | Disagree | -0.0476 | 0.0475 |
| | | 3-5 | Neither agree nor disagree, agree or strongly agree | -0.2537 | 0.0462 |