ONE APPROACH TO OVERSAMPLING BLACKS AND HISPANICS: THE NATIONAL ALCOHOL SURVEY

Robert L. Santos, Temple University

1. Introduction

The 1984 National Alcohol Survey (NAS) is the first national household survey to study exclusively the drinking attitudes, patterns and problems of Blacks, Hispanics and the general population. The survey objectives called for an area probability sample of households in the 48 coterminous states. In total, 5,000 one hour personal interviews were to be conducted with adults aged 18 or over. However, 1,500 interviews of Blacks, 1,500 interviews of Hispanics and 2,000 interviews from the remaining populations were desired.

The March 1984 Current Population Survey estimates that about 10.8 percent of all households in the U.S. are Black households, while only 5.1 percent represent Hispanic households. Thus, the problems of rare element sampling required special consideration. Kish (1965b) and Kalton and Anderson (1984) outline methods suitable for surveys of rare populations. Among such techniques, the NAS employed a multi-purpose sample (i.e., selection of Blacks, Hispanics and others), disproportionate stratified sampling of minority strata, large clusters, and supplementation (of primary sampling units).

The purpose of this paper is to detail the sampling design employed in the NAS. The results of the NAS are presented and a non-response analysis is conducted. The report closes with an examination of sampling error in the NAS. We shall show that the desired number of interviews was attained or exceeded for Blacks and Hispanics. However, large sampling errors resulted from differential weighting.

2. Selection Strategy

Several factors influenced the design strategy of the sample. First, the oversample of Blacks and especially Hispanics posed a formidable problem because of their rare nature. Differential sampling rates for each group would be required, and the implementation of such a design can be rather complex. Secondly, the entire budget available for design, data collection and data reduction was confined to just under $600,000. A per unit cost of $120 for this survey meant that only cost effective sampling plans could be considered. Finally, a third major constraint was time. According to contractual specifications, six months were allotted for sample design, selection, listing, selection of households and preparation of interviewer assignments.

In reaction to the conditions specified above the following design features were utilized in the NAS:

° The ISR 100 Primary Sampling Unit (PSU) National Sampling Frame was employed. This reduced both cost and time components in the first stage of a multi-stage design.

° Ten supplemental PSUs were selected to reduce variances for the Black and Hispanic portions of the survey. The NAS thus comprised a total of 110 PSUs.

° Secondary Sampling Units (SSUs) were stratified by density Hispanic and density Black population and heavier minority strata were oversampled. This reduced the screening costs of the survey.

° In the third stage of selection, households in minority strata were again oversampled. This further curtailed the costs of identifying Black and Hispanic households, and helped to equalize interviewer workloads and minimize intra-PSU travel costs.

° Small interpenetrating replicated "reserve" samples of households were set aside and later allocated in a fashion which ensured that desired subgroup sample sizes were achieved or exceeded.

Several features of the sample design of the NAS are somewhat similar to those utilized by Ericksen (1976). However, Ericksen's survey objectives did not include an Hispanic component.

3. Brief Description of the Sample Design

The NAS is based on a four-stage area probability sample of households in the 48 coterminous United States. In the first stage of selection, PSUs comprised groups of metropolitan counties representing Standard Consolidated Areas (SCAs) or Standard Metropolitan Statistical Areas (SMSAs), and groups of nonmetropolitan counties. Using 1985 projected population sizes, a total of 44 self-representing PSU equivalents were determined, while 66 noncertainty PSUs were selected with probabilities proportional to measures of size (pps). The final sample of 110 PSUs is essentially a 100 PSU national sampling frame (based on the general population) supplemented with 10 PSUs to capture the heterogeneity of the Black and Hispanic populations.

In the second stage of selections, Block Groups (BGs) and Enumeration Districts (EDs) defined SSUs. Within each PSU, SSUs were assembled into three distinct strata:

- an Hispanic stratum, consisting of all SSUs with 15 percent or more Spanish Origin population according to 1980 Census data;
- a Black stratum, consisting of all non-Hispanic SSUs with 10 percent or more Black population (in 1980);
- a "balance stratum" of all remaining SSUs.

The Hispanic stratum was oversampled by a factor of 8, while the Black stratum was selected at 3.56 times the rate of the balance stratum. Probabilities proportional (within a PSU) to 1980 population counts were used at this stage. In all, 581 SSUs comprised the sample.

Within each SSU, a single tertiary unit called a Listing Area (LA) was selected. A Listing Area is a group of one or more blocks which contain about 44 housing units (in 1980) on average. Enumerators were sent to each LA so that complete, up-to-date listings of housing units would be available for subselection.

In the final stage of selection, segments of about 4 contiguously listed housing units were sampled from listing sheets and assigned to interviewers for screening. Segments in Black and Hispanic LAs were selected at twice the rates of those in the balance stratum. In all, 2,844 segments were employed in the sample. A replicate reserve sample 25 percent the size of the initial sample was drawn in the event a larger sampling

fraction was desired. Near the end of data collection, the reserve sample was released into the field for all but the balance stratum. This ensured that desired minority subgroup sample sizes were attained or exceeded in the survey, but introduced larger differential weight factors.

In order to control sample sizes and reduce the effects of differential weighting, white (for ease of presentation, a "white" household actually denotes a non-Black, non-Hispanic household) households were subselected in Black and Hispanic LAs, and Black households were subsampled in Hispanic LAs. Subselection was achieved through the use of randomly assigned screening forms in predetermined proportions. Three screening form types were used. The first allowed an interview from a household regardless of the race/ethnicity of the household head; the second form permitted an interview only if the household head was reported as Black or Hispanic. The third form restricted interviews to households with Hispanic heads. These forms were randomly allocated in fixed proportions in a fashion which yielded (apart from the reserve sample allocation) an equal probability sample of white households, and an equal probability sample of Black households in all but the balance stratum.

4. Identification and Adaptation of the Sampling Frame of PSUs

The choice of a sampling frame was an important design issue in the NAS. Because the survey objectives called for a national area probability sample of households, the ISR 100 PSU National Sampling Frame was a natural choice. This frame is based on 1980 Census data and represents the population in the 48 coterminous states plus the District of Columbia. It is flexible enough to accomodate both general population surveys as well as studies of special subpopulations such as Blacks and Hispanics. The ISR National Sampling Frame was especially convenient because of the existing time restrictions and cost constraints.

The actual selection of NAS PSUs was composed of two distinct tasks:
° the selection of one hundred PSUs into the ISR national sample;
° the adaptation of the national frame through the selection of ten supplemental PSUs.

4.1 ISR National Sample of PSUs

Data from the 1970 and 1980 censuses were employed in generating 1985 population projections for all PSUs in the ISR national sample. The projections simply extrapolated the population changes which occurred in the 1970s:

MOS=1980 population + ½(1980 population – 1970 population).

These projections were used as measures of size in the selection of primary areas. These measures approximately equal 235 million when summed across all counties in the geographic area of coverage.

We defined PSUs as follows. First, counties were assembled into two groups: 1) those in "self-representing areas," and 2) the rest of the country. Self-representing areas are Standard Metropolitan Statistical Areas (SMSAs) or Standard Consolidated Areas (SCAs) with projected populations of two million or more. Eighteen self-representing areas satisfy this rule; they

contain a total projected population of 84.6 million, 36 percent of the national total.

For the rest of the country, PSUs were constructed from SMSAs and counties outside SMSAs in one of two ways: if an SMSA or county had a population of 150,000 or more, it was defined as a PSU; if an SMSA or county had a smaller population, it was combined with adjacent counties or a nearby SMSA to form PSUs with populations of 150,000 or more.

The next step involved the construction of thirty-two strata, each with total projected populations ranging 4.2 to 5.2 million. Collectively, these strata contained a projected population of 150.4 million. Strata were created with the goal of increasing their homogeneity. To this end, we employed region and metropolitan/nonmetropolitan status, and within these categories, one or more of the following variables were used: degree of urbanization, economic growth rates, racial composition, and the proportion of the population Hispanic.

Some strata included only a few metropolitan PSUs having populations of 500,000 or more. Other strata were rural in nature, with populations under 200,000 and with individual PSUs which included many counties covering a vast land area. In some cases, a single PSU covered over half the nonmetropolitan area of western or plains states.

Two PSUs were selected from each stratum; this yielded a total of 64 non-self-representing areas. PSUs were drawn with probabilities proportional to size. Thus, for PSU j in a given stratum,

$$f_{j*} = (2 \times MOS_{j*})/\text{stratum size},$$

where MOS_{j*} is the measure of size assigned to a given PSU j, and f_{j*} is the PSU selection rate.

The two selections were determined independently, or with replacement.

4.2 Supplemental PSU Selection

In addition to the 100 PSUs of the ISR National Sampling Frame, ten supplemental PSUs were employed. We supplemented the national frame in order to reduce the average cluster sizes of those PSUs with significant minority populations. Had no PSUs been added, unduly large numbers of interviews would have been taken from minority neighborhoods in non-self-representing PSUs. This would occur because the first stage selection probabilities were based on total population projections (rather than minority) and consequently were sometimes small. Moreover, substantial oversampling of minority strata was planned at subsequent stages. In order to yield an overall sampling rate f in the Hispanic stratum, for instance, a within PSU rate of f/p_i is necessary, where p_i is the PSU selection rate. If p_i is small then f/p_i could be large.

Through contractual agreements, four PSUs were targeted to increase the precision of the Black oversample, and six were designated for the Hispanic sample. Black and Hispanic supplemental PSUs were determined independently. Naturally, their selection necessitated changes in the structure and specification of the national sampling frame. These issues as well as the supplemental PSU selection procedure will now be discussed.

The self-representing PSUs in the ISR frame contain about 45 percent of the total Black population according to 1980 Census data. Roughly three quarters of the remaining Black population resides in the South. Thus, Black supplemental PSUs were confined to the Southern states. Three of thirteen Southern non-self-representing PSU strata in the ISR frame were constructed on the basis of high Black population density (25 percent or more). One stratum comprised rural PSUs while the remaining two contained metropolitan areas. Since the six PSUs representing these strata were most likely to experience large sampling rates within PSUs, additional PSU selections were drawn. One selection was drawn with pps from each metropolitan stratum, and two additional PSUs were drawn from the rural PSU stratum.

The advantages of this strategy were twofold. First, the selection of additional PSUs within the existing strata of the ISR frame was the most straightforward method of supplementing the sample. Secondly, time constraints did not permit more elaborate and perhaps more efficient supplementation strategies.

Hispanic supplemental PSUs were determined differently. About 88 percent of the Spanish Origin population resided within SMSAs in 1980. We therefore restricted PSU supplementation to metropolitan areas. We ranked all SCAs and SMSAs according to 1980 Spanish Origin population and designated those with Spanish Origin populations exceeding 150,000 as self-representing. All SCAs and two SMSAs were already self-representing and thus had no effect on the ISR frame. Three SMSAs had been selected into the ISR national sample as non-self-representing. Six previously unselected PSUs were added to the national sample as self-representing Hispanic supplements. In total, these self-representing areas contained roughly 66 percent of the total Spanish Origin population in the U.S.

The ISR national frame was altered to reflect the fact that nonself-representing PSUs were now self-representing. The net effect on the ISR frame was that nine PSUs were deleted from five strata. Stratum population totals were adjusted to their new sizes, and in one case, two strata were collapsed to preserve the paired selections design.

5. Second and Third Stages of Selection

The second stage of selection was accomplished in three steps: (1) creation of SSUs; (2) stratification of SSUs; and (3) selection of SSUs. SSUs were defined as EDs or BGs which contained a minimum of 44 housing units in 1980. Whenever an ED or BG did not meet the minimum, it was combined with neighboring units until this criterion was satisfied. Within each PSU, SSUs were assembled into three strata:

- an Hispanic stratum, consisting of all SSUs with 15 percent or more Spanish Origin population in 1980;
- a Black stratum, composed of all non-Hispanic SSUs with 10 percent or more Black population;
- a balance stratum of all remaining SSUs.

The Black and Hispanic strata were oversampled at this stage of selection in order to substantially reduce screening costs. Relative to the balance stratum, Hispanic SSUs were oversampled by a factor of 8, while Black SSUs were oversampled by 3.56. An important design consideration is the efficiency of oversampling. The efficiency depends on the density of minorities (Black and Hispanic) within the minority SSU strata, as well as the proportion of the minority populations in those strata. A full discussion of this topic is furnished in Waksberg (1973). At the time the sampling rates were set, however, these data were not available.

A total of 581 SSUs were drawn, of which 188 were Hispanic, 179 were Black and 214 were from the balance stratum. A larger number of SSUs was drawn in order to curb the ill effects of intraclass correlations associated with sampling within PSUs. One drawback is that a large number of SSUs increases intra-PSU travel. This lessens the savings from reduced screening costs.

In the third stage of selection, a single sampling unit called a Listing Area was selected from each SSU. The selection was made with pps using 1980 Census data and the SSU selection probability. Each LA contained about 44 housing units on average. The resultant sample of LAs comprised an equal probability sample within the Black, Hispanic and balance strata. The overall selection equation may be written as follows: For an LA in nonself-representing PSU i, in SSU stratum j

$$(2)\ f_j = \frac{PSU_i}{Stratum\ total} \times \frac{(K_j)NUMLA_j}{TOTLA_i} \times \frac{1}{NUMLA}$$

where

stratum total = 1985 projected population total in that PSU stratum

PSU_i = 1985 projected population total for PSU_i

$$K_j = \begin{cases} 8, & \text{for } j = 1,\ \text{the Hispanic stratum} \\ 3.56, & \text{for } j = 2,\ \text{the Black stratum} \\ 1, & \text{for } j = 3,\ \text{the balance stratum} \end{cases}$$

$TOTLA_i$ = (12,800) PSU_i/(Stratum total)

NUMLA = $TOTLA_i$ × (1980 population in SSU)/ (1980 population in PSU i)

$$f_j = \begin{cases} 1/800 & \text{for } j = 1 \\ 1/1,800 & \text{for } j = 2 \\ 1/6,400 & \text{for } j = 3. \end{cases}$$

For self-representing PSUs, the "first" stage selection probability is one, and the "second" stage fraction becomes (K_j)(NUMLA)/6,400.

6. Household and Respondent Selection

Households were selected using a two phase methodology. In the first phase, segments of about 4 households were sampled at a constant rate of 6/10 for the Black and Hispanic strata, and 3/10 for the balance stratum. A total of 2,844 segments were employed in the sample. Oversampling by a factor of two helped to equalize interviewer workloads.

Three quarters of selected households were randomly designated for screening and interviewing. The remainder was split into two equal sized replicates and set aside as a reserve sample. It would be used to boost the sampling fraction and meet prespecified numbers of interviews. Midway through the data collection period, we decided to allocate only the reserve samples in the Black and Hispanic strata. This was done in reaction to a lower than expected number of Black and Hispanic interviews gathered by that time. While ensuring

that desired numbers of interviews were achieved or exceeded, the reserve allocation yielded an additional differential weighting factor of 4/3 between the minority and balance strata. At this point, the overall household selection probabilities were 1/1,333, 1/3,000, and 1/28,444 for the Hispanic, Black and balance strata, respectively.

The second phase of household selection consisted of randomly subsampling white households in the Hispanic and Black strata, and subsampling Black households in the Hispanic stratum. The objective was to subsample certain households in minority strata in a fashion which yielded an equal probability sample of white households overall, and an equal probability sample of Black households in the Black and Hispanic strata. To this end, three color coded screening forms were allocated randomly in fixed proportions to all sample households. Yellow, blue and ivory screening forms were used. A yellow form permitted an interview from a randomly selected adult resident if the household head was reported to be Hispanic in origin; otherwise, no interview was conducted at that household. With blue forms, an interview was conducted if the household head was either Black or Hispanic. Ivory forms allowed an interview regardless of the race/ethnicity of the household head.

Within a given SSU stratum, screening forms were distributed to sample households in the percentages shown in Table 1. Apart from the differential weighting due to the allocation (which was unanticipated) of the reserve sample, the form distributions yielded an equal probability sample of white households overall, and of Black households in minority SSU strata. The final household probabilities of selection are presented in Table 2. Note that the selection probability is contingent upon the race/ethnicity of the household head and the SSU stratum. The fractions in Table 2 also incorporate the increased selection probabilities in the minority SSU strata due to the reserve sample allocation.

One adult aged 18 or over was randomly selected within each eligible household. The random selection was realized using standard selection tables outlined in Kish (1949).

6. Survey Results

A total of 10,925 households were selected in the NAS. Table 3 presents screening response rates by SSU strata. Overall, about 92 percent of all sample households were screened. Screening response rates varied only slightly across SSU strata, ranging 89 percent in the balance stratum to 93 percent in the Hispanic stratum.

Final dispositions of screened households appear in Table 4. The first row of this table shows that about one third of white households were subselected out of the sample in the Black stratum, and roughly half of non-Hispanic households were subsampled in the Hispanic sample. (When a household in the Black or Hispanic stratum was subselected out of the survey on account of the race-ethnicity of the household head, the final disposition was called a designated termination.) The first two rows of Table 4 denote the successful completion of an interviewer's work. As such, work on about 88 percent of all screened households was completed.

Interview response rates must be based on the last two rows of Table 4. The interview response rates were 83 percent for the Black stratum, 77 percent for the Hispanic stratum and 82 percent in the balance stratum. Overall, an interview response rate of 81 percent was obtained. It is interesting to note that the Hispanic stratum displayed the highest screening response and the lowest interview response, while the screening and interview rates within the Black and balance strata were roughly the same.

Total estimates of survey response may be obtained by multiplying the screening and interview response rates. Doing this, we see that the overall response rate in the Black stratum was 75.9 percent; the overall response rate for the Hispanic stratum was 72.2 percent; and the overall response rate in the balance stratum was 73.2 percent. The total survey attained an overall response rate of 74.2 percent.

In all, 5,221 interviews were conducted in the NAS. Blacks accounted for 1,947 interviews, Hispanics totaled 1,433 interviews and 1,841 cases represent all others. The distribution of interviews by race/ethnicity and SSU stratum is produced in Table 5. The results of oversampling are clearly revealed in this table. Ninety-two percent of all Black interviews were taken from the Black stratum; similarly, 93 percent of Hispanic interviews were drawn from the Hispanic stratum. The non-Black, non-Hispanic interviews are more evenly spread across SSU strata, although the balance stratum produced 82 percent of these cases.

The results of Table 5 can perhaps be put into better perspective by noting the distribution and density of the minority and other populations across the SSU strata. Estimates based on 1980 Census data have been produced for the metropolitan (SMSA) areas of the U.S. and are presented in Tables 6A and 6B. The first row of Table 6A suggests that the SSU stratification was quite successful in isolating 81 percent of the Black population in the Black stratum while retaining only 18 percent of the total population. This gives rise to a 56.5 percent Black population density, as evidence in the first row of Table 6B.

The SSU stratification scheme was not as successful in isolating the Hispanic population. In the second row of Table 6A, we see that 63 percent of the Hispanic population is contained in the Hispanic stratum. The density of Hispanics in this stratum was almost 50 percent (see Table 6B). Unfortunately, 28.5 percent of Hispanics remained in the balance stratum. Because this stratum was substantially undersampled, we should expect large sampling errors and design effects for Hispanics.

Average design effects (DEFF), square roots of design effects (DEFT) and intraclass correlation coefficients (ROH) are presented for selected subgroups in Tables 7 and 8. The design effect is defined as the ratio of the variance of a statistic obtained from a complex sample design to that obtained from a simple random sample of the same size. The intraclass correlation is approximated by the following:

$$ROH \doteq (DEFF-1)/(\bar{b}-1)$$

where DEFF denotes the design effect and $\bar{b}$ is the average number of interviews per PSU. A discussion of these statistics may be found in Kish (1965a).

The average DEFFs, DEFTs and ROHs in Tables 7

and 8 are based on estimated proportions for 18 questionnaire items relating to alcohol consumption and abuse. Table 7 is based on the unweighted data set, while Table 8 was obtained using weighted data. The average DEFFs in Table 7 range from 1.57 for males to 2.17 for the total sample. Average ROHs range .025 for the total sample to .067 for non-Blacks/non-Hispanics. Given the substantial efforts to reduce intraclass correlations (e.g., PSU supplementation, many SSU selections), the DEFFs and ROHs most likely reflect the natural intraclass correlations associated with measurement of alcohol related items in a cluster sample.

Table 8 shows the impact of oversampling on the variances of sample proportions. Average design effects at least doubled for the total sample and sex subgroups. Standard errors for these groups increased by roughly 50 percent. For the Hispanic and Mexican subgroups, DEFFs based on weighted data are about triple the values obtained from the unweighted data. This represents an increase in standard error or over 80 percent. Design effects for Blacks and Others increased the least. For Blacks the increase in standard error was 29 percent, while for the subgroup of other races-ethnicities, the increase was 21 percent.

7. Conclusions

The design and results of the National Alcohol Survey are useful in illustrating several points regarding rare element sample surveys. First, the economic design and collection of survey data for Black and Hispanic populations is possible, especially in surveys which require Black, Hispanic and "other" components. Statistically efficient designs for Black oversamples are easier to attain than Hispanic oversamples because the Black population concentrates itself more in higher density minority areas than Hispanics do.

Adequate response rates can be obtained in minority oversamples. The NAS screening response rate was 92 percent and the interview response rate was 81 percent.

Disproportionate stratified sampling is the key to economic and efficient surveys. With regard to the NAS, the Hispanic stratum was oversampled by too large a factor. This was the price paid to achieve the survey goal of about 1,500 Hispanic interviews. In retrospect, it would have been more efficient to not allocate the reserve sample, accept about 400 fewer Hispanic interviews and incur smaller weighting effects.

It is possible that a more conservative definition of the Hispanic stratum could have permitted a more efficient design. For instance, one might employ a minimum criterion of 5 percent or more 1980 Hispanic population within an SSU (instead of 15 percent). This could have boosted the percentage of the Hispanic population within the Hispanic stratum substantially. However, this would also lessen the density of Hispanics and therefore increase the per unit cost of an Hispanic interview in that stratum.

REFERENCES

Description of the 1980 ISR National Sampling Frame. Institute for Survey Research, Temple University, Philadelphia, Pennsylvania.

Ericksen, E. (1976) Sampling a rare population: a case study. Journal of the American Statistical Association, Vol. 71, pp. 836–822.

Kalton, G. and Anderson, D. (1984) Sampling rare populations for health surveys. Proceedings of the Section on Survey Research Methods, American Statistical Association.

Kish, L. (1949) A procedure for objective respondent selection within the household. Journal of the American Statistical Association, Vol. 44, pp. 380–387.

Kish, L. (1965a) Survey Sampling. Wiley, New York

Kish, L. (1965b) Selection techniques for rare traits. In Genetics and the Epidemiology of Chronic Diseases, Public Health Service Publication No. 1163.

Waksberg, J. (1973) The effect of stratification with differential sampling rates on attributes of subsets of the population. Proceedings of the Social Statistics Section, American Statistical Association.

Table 1: Percentagewise Distribution of Screening Forms by SSU Stratum for the National Alcohol Survey

	Color of Form		
SSU Stratum:	Yellow	Blue	Ivory
Hispanic	55.56%	38.19%	6.25%
Black	–	85.94%	14.06%
Balance	–	–	100 %

Table 2: Household Selection Probabilities in the NAS by SSU Stratum and Race/Ethnicity of Household Head

	Race/Ethnicity of Household Head		
SSU Stratum:	Hispanic	Black	Other
Hispanic	1/1,333	1/3,000	1/21,333
Black	1/3,000	1/3,000	1/21,333
Balance	1/28,444	1/28,444	1/28,444

Table 3: Screening* Response Rates in the National Alcohol Survey by SSU Stratum

Screening Disposition:	SSU STRATUM			Total
	Black	Hispanic	Balance	
HH Not Screened	378	302	235	915
(column %)	(9.0%)	(6.6%)	(10.9%)	(8.4%)
HH Screened	3,830	4,268	1,912	10,010
(column %)	(91.0%)	(93.4%)	(89.1%)	(91.6%)
Total	4,208	4,570	2,147	10,925
(column %)	100%	100%	100%	100%

*This table excludes 1,191 selections which were found not to be in the sample universe (e.g., outside sample area, vacant, dilapidated, business, etc.).

Table 4: Final Disposition of Screened Households
by SSU Stratum in the NAS

| Final Disposition: | SSU STRATUM | | | |
	Black	Hispanic	Balance	Total
Designated Termination*	1,388	2,180	–	3,568
(column%)	(36.2%)	(51.1%)		(35.6%)
Respondent Interviewed	2,036	1,615	1,570	5,221
(column %)	(53.2%)	(37.8%)	(82.1%)	(52.2%)
HH Member Not Interviewed	406	473	342	1,221
(column %)	(10.6%)	(11.1%)	(17.9%)	(12.2%)
Total HHs	3,830	4,268	1,912	10,010
(column %)	(100%)	(100%)	(100%)	(100%)

*Designated Termination denotes those households which were sub-selected out of the sample on the basis of the race/ethnicity of the household head.

Table 5: Distribution of Interviews by Race/Ethnicity
of Respondent and SSU Stratum in the NAS

| Race/Ethnicity: | SSU STRATUM | | | |
	Black	Hispanic	Balance	Total
Blacks	1,785	135	27	1,947
(row%)	(91.7%)	(6.9%)	(1.4%)	(100%)
Hispanics	62	1,338	33	1,433
(row%)	(4.3%)	(93.4%)	(2.3%)	(100%)
Others	189	142	1,510	1,841
(row%)	(10.3%)	(7.7%)	(82.0%)	(100%)
Total	2,036	1,615	1,570	5,221
(row%)	(39.0%)	(30.9%)	(30.1%)	(100%)

Table 6A: Percentage of the Black, Hispanic
and Total Populations in Each SSU Stratum
for the Metropolitan Areas of the U.S.

| Population: | SSU STRATUM | | | |
	Black	Hispanic	Balance	Total
Black	81.0%	9.3%	9.8%	100%
Hispanic	8.6%	62.8%	28.5%	100%
Total	18.1%	9.8%	72.1%	100%

Table 6B: Density of Black and Hispanic Population
by SSU Stratum

| | SSU STRATUM | | | |
	Black	Hispanic	Balance	Total
Density of Black Pop.	56.5%	12.0%	1.7%	14.0%
Density of Hispanic Pop.	3.6%	48.6%	3.0%	7.7%

Table 7: Average DEFFs, DEFTs and ROHs Based on Unweighted
NAS Data for the Total Sample and Selected Subgroups

	SAMPLE SIZE	AVE DEFF	AVE DEFT	AVE* ROH
Total Sample	5,221	2.172	1.460	.025
Sex:				
Males	2,093	1.571	1.247	.032
Females	3,128	1.940	1.373	.034
Race/Ethnicity:				
Blacks	1,947	1.758	1.315	.035
Hispanics	1,433	2.056	1.392	.041
Mexicans	949	1.862	1.809	.051
Other Race/Ethnicity	1,841	2.065	1.409	.067

*Average values of ROH were obtained using $(AVE.DEFF-1)/(\bar{b}-1)$ where $\bar{b}$ is the average number of interviews per PSU. The total number of PSUs in $\bar{b}$ is equal to the number of PSUs with nonzero interviews for that subgroup: 86 for Blacks; 53 for Hispanics and Mexicans; and 110 for all other subgroups.

Table 8: Average* DEFFs, DEFTs and ROHs Based on Weighted
NAS Data for the Total Sample and Selected Subgroups

	SAMPLE SIZE	AVE DEFF	AVE DEFT	AVE* ROH
Subgroup: Total Sample	5,221	4.966	2.194	.085
Sex:				
Males	2,093	3.351	1.813	.130
Females	3,128	4.307	2.048	.121
Race/Ethnicity:				
Blacks	1,947	2.936	1.698	.089
Hispanics	1,433	6.784	2.533	.222
Mexicans	949	5.888	2.386	.289
Other Race/Ethnicity	1,847	2.584	1.576	.100

*The average values of ROH were obtained using $(AVE.DEFF-1)/(\bar{b}-1)$ where $\bar{b}$ is the average number of interviews per PSU. The total number of PSUs in $\bar{b}$ is equal to the number of PSUs with nonzero interviews for that subgroup: 86 for Blacks; 53 for Hispanics and Mexicans; and 110 for all other subgroups.