USING THE CENSUS 2000/2001 SUPPLEMENTARY SURVEY AS A SAMPLING FRAME FOR THE NATIONAL EPIDEMIOLOGICAL SURVEY ON ALCOHOL AND RELATED CONDITIONS

Marie Stetser, Jana Shepherd, and Thomas F. Moore[1]
U.S. Census Bureau, Washington DC 20233-8700

KEY WORDS: Sampling frame, sample design, C2SS, double sampling

I. Introduction

In 2001 the U.S. Census Bureau launched field operations for the National Epidemiological Survey on Alcohol and Related Conditions (NESARC). NESARC is a survey sponsored by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and is designed to collect data on the prevalence of alcohol and drug abuse and associated psychiatric conditions in the United States. NESARC became the first survey to use the Census 2000/2001 Supplementary Survey (C2SS) as a sampling frame. In this paper we will discuss the sampling options we considered for NESARC and the decisions we made about the sample design. We will describe the design of the C2SS and how we used it in the design of NESARC. We will then evaluate the design and show what we learned about the effectiveness of oversampling in the design and the effect of the C2SS race and mode on response outcomes.

II. Options for NESARC

The target population for NESARC includes civilian non-institutionalized persons age 18 and over in all 50 states and the District of Columbia. NIAAA wanted a multi-stage, stratified sample design of about 48,000 completed responses that would include an oversampling of 9,000 Blacks and 9,000 Hispanics. NIAAA indicated they would like to use the same design we used for a survey we conducted for them in 1992 called the National Longitudinal Alcohol Epidemiological Survey (NLAES). For NLAES we used extra sample from the National Health Interview Survey (NHIS). Unfortunately, this time there was no extra sample available from NHIS, so we had to look for other alternatives to obtain a sample.

The first option we considered was to use extra sample from the National Crime Victimization Survey (NCVS). This option would give NESARC a multi-stage, stratified design with 245 primary sampling units (PSUs), 93 of which were self-representing (SR) and 152 of which were nonself-representing (NSR). At the second stage of sample selection, housing units were selected in clusters of four adjacent units. For this option, we would have no information about the household demographics of any of the sample cases prior to the NESARC interview. In order to obtain the desired number of minority responses, we would have had to inflate the sample size and then screen out some non-Black, non-Hispanic households to reduce field costs.

The second option we considered was to use the C2SS as a sampling frame. The C2SS is a large intercensal rolling survey that uses the American Community Survey (ACS) questionnaire. The C2SS was used to test the feasibility of the ACS, which will begin in 2003, and may one day replace the long form in the Decennial Census. For NESARC we proposed using the respondent information collected in the C2SS to target Black and Hispanic households. The C2SS would also give NESARC a multi-stage, stratified design, but had 655 PSUs, 366 of which were SR and 289 of which were NSR. There was no clustering of housing units in the second stage.

When we compared the two options, we first looked at how many interview households we would need to attempt to interview in order to accomplish our sponsor's goals for oversampling. In order to calculate a sample size for the NCVS option, let's assume the proportion of Blacks in the United States to be 12.3% and the proportion of Hispanics to be 12.5% (U.S. Census Bureau, *Statistical Abstract* 25, 27). If we wanted the completed interviews to include 9,000 Blacks and 9,000 Hispanics, we would need to screen at least 73,170 households. When we factor in a generous 92% response rate, a 4.7% rate of permanently ineligible cases (demolished, businesses, etc), and a 13.5% rate of temporarily ineligible cases (vacants, etc), we get a starting sample size of over 96,000 addresses.

For the C2SS option, the oversampling did not drive the sample sizes as much. To use the C2SS option, we decided that we would sort the C2SS respondent households into 3 categories: households with at least one Hispanic person, households with at least one Black person, and everything else (Black Hispanics were assigned to the Hispanic group).

[1]This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

Since for NESARC we would be interviewing only one person per household, we looked at the race and ethnicity distributions within the minority households from the January 2000 panel and estimated the average probability of selecting a Hispanic person or a Black person from households that were labeled respectively. For Hispanic households we estimated we would select a Hispanic 84.4% of the time, and for Black households we estimated we would select a Black 95.8% of the time. We assumed the same 92% response rate, but because we had already contacted these households in the C2SS, we assumed only a 1% rate for permanently ineligible cases and a 6.4% rate for temporarily ineligible cases. Using these assumptions we estimated that in order to meet our minority oversampling requirements we would need to sample 12,604 cases from the Hispanic stratum, 11,109 cases from the Black stratum, and 34,687 cases from the Other stratum for a total starting sample size of 58,400.

Comparing the sample sizes for the two options, we saw an opportunity to greatly reduce costs and response burden. By using the C2SS we avoided having to contact over 37,600 cases to conduct a costly screener interview. We did consider possibly using retired cases from NCVS and using the race and ethnicity information from the NCVS interviews. However, NIAAA felt strongly that this would cause a response problem because those households would have already been interviewed several times for NCVS before being interviewed for NESARC. Also, the C2SS option offered 410 more PSUs than the NCVS option and did not have the clustering of housing units in the 2nd stage of sampling. While there would be some additional field costs associated having more PSUs and more scattered housing units, we felt these would be unlikely to come close to the cost of having to do so many screener interviews. And as a bonus, there would be an increase in the precision of the estimates with the added PSUs and the absence of the 2nd stage clustering. Still, the overwhelming reason that NIAAA decided to use the C2SS option was to avoid having to use a screening interview to obtain the desired number of minority sample cases.

III. The C2SS Sample Design

In order to explain the design of NESARC we must first discuss the design of the C2SS. The C2SS covers the population in all 50 states and the District of Columbia with the exclusion of persons living in group quarters. The C2SS collects information about persons and housing units.

For the most part, the C2SS PSU definitions are the same as the county-based PSUs used in the Current Population Survey (U.S. Census Bureau, Bureau of Labor Statistics 3-3) with a few modifications to account for changes in MSA and county definitions. The universe of C2SS PSUs consists of all 3,142 counties and county-equivalents in the United States. The C2SS sample PSUs include 42 counties selected with certainty to serve as comparison counties with the 2000 Decennial Census. The remaining PSUs were included in the selection of the "National Sample." For the National Sample, PSUs with a 1996 population of 250,000 or more were selected as self-representing (SR). All other PSUs were designated as nonself-representing (NSR) and stratified within state by several demographic characteristics including population and housing unit growth, education, poverty, number of rented housing units, rural population, and in some states, Hispanic and Black populations. From each stratum, two NSR PSUs were selected with probability proportional to the size of the estimated 1996 population.

Once the sample PSUs were determined, a housing unit frame was constructed from the August 2000 version of the Master Address File (MAF). For the 2001 C2SS unit frame sample, we extracted eligible housing units from the MAF within each sample PSU. Before sampling we performed a detailed sort of the housing units by geography. The unit frame within-PSU sampling for each year occurred in 2 stages. In the first stage, we selected 17.5% of the housing units. In the second stage, we chose a systematic sample of units within each county using a specified sampling interval to obtain the required sample size for each PSU. In certain rural Census 2000 enumeration areas we included an area frame; about 66 counties included blocks where we performed address listing and sampling procedures.

For C2SS data collection, Census mailed the C2SS questionnaire to all sample housing units (HUs) with mailable addresses. For HUs that did not return the questionnaire in their panel month, we attempted to locate a telephone number and interview the respondents using computer assisted telephone interviewing (CATI). For HUs where we still hadn't obtained a response within two months we subsampled at a rate of 1 in 3 and attempted to interview the selected HUs using computer-assisted personal interviewing (CAPI). HUs that did not have a mailable address went directly to CAPI and were subsampled at a rate of 2 in 3.

IV. The NESARC Sample Design

We selected most of the NESARC sample from five panels of the C2SS: November 2000 through March 2001. After the interviewed cases from these five C2SS panels were completed, we determined which

cases to include for NESARC sample selection. We included vacant and occupied, non-seasonal housing units that responded to the C2SS. We excluded units identified as seasonal or where the household refused the C2SS CAPI interview. We also excluded units that were eligible for CAPI subsampling, but were not selected.

As previously described in section II, we used the information on race and ethnicity collected from the C2SS to sort the cases within each sample PSU into three substrata-- Hispanic, Black, and Other (Non-Black, Non-Hispanic). Units with no race or ethnicity information from the C2SS were included in the Other substratum. Because final weights for the C2SS data were not available at the time of sampling, we constructed representative weights that take into account the following: PSU selection probabilities, within-PSU selection probabilities, CAPI subsampling probabilities, and an adjustment for nonresponse in the C2SS. Because of the CAPI subsampling factor, there was extra variability among the weights with cases collected via CAPI having 1.5 or 3 times the weights of the mail and CATI cases. So to smooth out the weights and thereby, decrease the sampling variability, we selected the housing units with a probability proportional to their representative weight.

NIAAA also wanted to include persons in non-institutionalized, civilian group quarters for NESARC. Unfortunately, the C2SS did not cover group quarters, so we used as an alternative the Census 2000 Group Quarters Inventory to create a group quarters' frame. We converted the number of people from the 2000 Census living in each group quarters' unit to a housing unit equivalent number that we used to create a housing unit level weight. This allowed us to give persons living in group quarters similar probabilities of selection to those of housing units. Group quarters were sampled together with housing units in the Other substratum.

We selected the number of sample cases from each substratum as we planned in section II for the initial sample. We also selected some additional sample cases, which we held in reserve; several months into the data collection we released an additional 926 cases for interviewing, bringing our sample size to 59,326 units.

For each housing unit, field representatives listed persons 18 years and older and randomly selected one person from the roster. Persons age 18-24 were selected with a probability 2.25 times that of persons age 25 and older so as to create an over-sampling of young adults.

Data collection began August 15, 2001 and continued through April 2002. The overall weighted response rate was 88.2%[2].

V. Evaluating the NESARC Design - Analysis and Results

Because this is the first time Census has designed a survey using the C2SS, we wanted to evaluate the effectiveness of our sample design in achieving our oversampling goals. First we looked at the actual frequencies and percentages of minority cases we achieved overall and for each substratum. Table 1 shows the actual race and ethnicity outcomes of the NESARC sample by substratum.

Table 1. NESARC Race/Ethnicity Outcome by Substratum

Substratum		Race/Ethnicity of NESARC Respondent			
		Hispanic	Non-Hispanic Black	Non-Hispanic, Non-Black	Total
Hispanic	Number	7,493	237	2,108	9,838
	Percent	76.2	2.4	21.4	100.0
Black	Number	193	7,585	737	8,515
	Percent	2.3	89.1	8.7	100.0
Other	Number	708	604	24,152	25,464
	Percent	2.8	2.4	94.8	100.0
Total	Number	8,394	8,426	26,977	43,817
	Percent	19.2	19.2	61.6	100.0

Note: There were 3,290 cases with missing race and ethnicity data.

When we were designing the NESARC sample, we knew that the distributions would not turn out exactly as planned. We knew there was a 6-10 month lag (depending on which C2SS panel the household was interviewed in) between the time the C2SS responses were collected and when we fielded the NESARC survey. We fully expected that people would move, households would change their demographic make-up, and there would be some deterioration of the assumptions. However, our objective was to merely increase the number of Hispanic and Black persons selected in the NESARC sample – not to necessarily contact the exact same people that responded to the C2SS. We believed that the design would be forgiving. We knew in some cases where we were expecting Blacks or Hispanics, we would get non-Black or non-Hispanic respondents and vice versa. Even though we had no data that could help us predict exactly how much the race and ethnicity assumptions would deteriorate, we speculated that these would end up being balanced over the sample.

Because of the lower than expected response rate

[2]Response rate was calcuated as NESARC completed interviews divided by eligible sample units. This figure does not take into account the nonresponse from the C2SS.

we felt it was better to analyze the percentages rather than the sample size counts. For the Hispanic substratum, 76.2% of the respondents reported they were Hispanic; for the Black substratum 89.1% of the respondents reported they were Black. Fortunately, our suspicions were correct. As the table shows, we picked up enough minority cases in the alternate substrata to make up the differences and then some. Overall, the respondents included 19.2% Hispanics and 19.2% Blacks. This is slightly higher than the 18.75% that we were expecting in each group. So overall, our oversampling was very successful.

When we were designing the sample we were concerned that the race and ethnicity data would deteriorate over time; we hypothesized that the more time that had passed between the C2SS interview and the NESARC interview, the less effective the oversampling would be. Table 2 shows the percentage of respondents where the actual race or ethnicity matched that of the substratum by the C2SS panel where the household was originally sampled. Surprisingly, for all three substrata, there were no significant differences among the C2SS panels. Table 3 shows the percentage of respondents where the actual race or ethnicity matched that of the substratum by the number of months between the C2SS interview and the NESARC interview. For the Black and Other substrata, there was only a slightly smaller percentage of matches for respondents where there were 11-17 months between the C2SS response and the NESARC interview, compared with those where there were only 6-8 months and 9-10 months between interviews. For the Hispanic substratum, there were no significant differences among the groups.

We also had some questions as to whether we should use different response assumptions for each substratum. Were households in the minority substrata more or less likely to respond than those in the non-minority substrata? Would the percentages of temporarily or permanently ineligible cases differ among the substrata? Table 4 shows the Response outcomes by substratum. For each substratum, the percent of completed interviews was significantly different from the other two substrata. The Hispanic substratum had the highest percentage of completed interviews at 81.8%, followed closely by the Black substratum at 80.1%. The Other substratum had only 77.6% completed interviews (not shown). We also noticed higher rates of temporarily and permanently ineligibles in the Other substratum. So to further explain this, we broke out the Other substratum into three categories: units with a valid Non-Black, Non-Hispanic race response in C2SS, units that were vacant or had a missing race response in C2SS, and units that came from the group quarters frame. Interestingly, the percentage of eligible non-respondents did not differ significantly among the substrata or the 3 Other categories. However, the percentage of temporarily ineligible did differ among each group. The Non-Hispanic, Non-Black with valid race responses had the lowest percentage of temporarily ineligible at 6.5%, followed closely by the Hispanic substratum at 7.7% and the Black substratum at 9.0%. The group quarters and cases with missing race/vacant group had much higher rates of temporarily ineligible at 17.4% and 38.7%. Also, the percentage of permanently ineligible cases is significantly higher for the these two groups at 15.6% for group quarters and 5.7% missing race/vacant, compared with 0.8% and 0.7% for the Hispanic and Black substrata[3]. Finally, looking at the overall response rates among the categories, we were not surprised to see that the Other cases that were vacant or had a missing race response did have a lower overall response rate than the 3 groups where there was a valid race response. Based on these outcomes, we recommend for future designs that the "Other" group be further stratified in future to improve the efficiency of the sampling.

Finally, we looked at the effect that the C2SS mode had on the NESARC response outcome (Table 5). The housing units that were previously interviewed in the C2SS via CAPI had a much smaller percentage of NESARC completed interviews (68.5%) compared to housing units that completed the C2SS by mail (84.8%) or by CATI (82.7%). However, the response rate is a more accurate way to measure the percentage of completed interviews because it only includes the eligible cases in the calculation. Looking at these numbers, the housing units previously interviewed through CAPI still had a lower rate (84.8%) than those interviewed by CATI (88.5%) or by mail (90.5%) but not by as much. The percentage of eligible NESARC nonrespondents increased with each mode of data collection with those previously interviewed by CAPI having the highest rate of eligible nonrespondents and those that responded by mail having the lowest. It's not surprising that the reluctant responders to the C2SS would also be less likely to respond to NESARC. Mail cases had the highest propensity to respond because they responded soonest to the C2SS. Housing units that we contacted by CATI had been somewhat more reluctant to respond, and housing units that we resorted to interviewing via CAPI were perhaps the most reluctant. For the temporarily ineligible group and the permanently ineligible group, respondents who

[3] The Black and Hispanic permanently ineligible percentages are not significantly different from each other.

responded via CAPI to the C2SS again had significantly higher percentages than those that responded via CATI or mail. For future research, we could consider comparing the detailed outcomes for both the C2SS and NESARC to further explain this.

VI. Summary and Conclusions

So finally, to summarize what we've learned:

Our design of using the C2SS as a sampling frame for NESARC and oversampling minorities was very effective in obtaining the desired sample distribution of Blacks and Hispanics.

There is evidence to suggest that the race information deteriorated the longer the time between the C2SS interview and the NESARC interview. However, the time lapse did not hinder our ability to obtain the desired number of minority respondents.

The percentage of completed interviews varied somewhat by the race/ethnicity substratum. The Other substratum seemed to have the largest difference, particularly in the percentages of ineligible units. This was further explained by expanding this substratum into 3 groups, comparing units where there was a valid race response in C2SS with those that were vacant or had missing race response and the group quarters units. In the future we recommend sampling each of those 3 groups in separate strata.

The mode of the C2SS interview did affect the likelihood of obtaining a completed NESARC response. C2SS respondents who were interviewed via CAPI had the highest proportion of eligible nonrespondents, temporarily ineligible, and permanently ineligible cases.

The primary advantage of using the C2SS as a sampling frame is that we can successfully target subpopulations of interest without having to contact large numbers of households just to find that population. The C2SS is a vast resource of information that we could use to design sample surveys in the future. The C2SS could be used for surveys that, in the past, were not feasible because it was very difficult to create a frame of the target population and screening procedures would have been very expensive. For NESARC we targeted the Black and Hispanic populations, but for another survey, we may choose to target something more specific, such as households with children or households below a certain income level. For those types of characteristics, the data may be less forgiving and may deteriorate faster; the lag between when the C2SS data are collected and when the survey is fielded may have a greater effect on locating the target population than it did in NESARC. However, we could likely mitigate this by reducing the amount of time between when the C2SS data is collected and when the data collection for the survey begins. Based on our overall experience with NESARC, we would feel comfortable using this sample design again.

As a final note, concerns have been expressed about whether the use of C2SS or ACS as a sampling frame would give the Census Bureau an unfair advantage over private sector survey firms. The Census Bureau is not currently working on any additional surveys using the C2SS or the ACS as a sampling frame. The Bureau will work closely with the Office of Management and Budget to develop a clear policy outlining the circumstances and terms under which it is permissible to use the ACS as a sampling frame. We recognize that this policy must avoid harm to the private sector while still meeting the confidentiality requirements of Title 13 and meeting our obligations to the taxpayers.

Bibliography

National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health. "Strategic Plan To Address Health Disparities." February 8, 2001. http://www.niaaa.nih.gov/about/Disparities01.htm

National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health. "Alcohol Epidemiologic Directory" June 2001. http://www.niaaa.nih.gov/publications/2001%20AEDS%20Data%20Directory.pdf

U.S. Census Bureau, *Statistical Abstract of the United States:2001* (121st edition), Washington D.C. 2001.

U.S. Census Bureau, Demographic Statistical Methods Division, "Memorandum from Cahoon to Lewis: National Epidemiological Survey of Alcohol and Related Conditions (NESARC) Universe Creation and Sampling Specifications."

U.S. Census Bureau, Bureau of Labor Statistics, *Current Population Survey: Design and Methodology*, Technical Paper 63RV, Washington, DC, 2002 http://www.census.gov/apsd/techdoc/cps/cps-main.html.

Table 2. Percent of Respondents Where the Actual Race/Ethnicity Outcome Matched the Substratum by C2SS Panel[4]

Race/Ethnicity Group		C2SS Panel				Total of All Panels
		November 2000	December 2000	January 2001	February/ March 2001	
Hispanic	Number	1726	1842	1897	2028	7493
	Percent (SE[5])	76.5 (0.99)	77.1 (0.96)	78.4 (0.93)	78.7 (0.89)	76.2
Black	Number	1838	1821	1869	2057	7585
	Percent (SE)	88.8 (0.73)	90.1 (0.69)	88.5 (0.72)	88.5 (0.69)	89.1
Other	Number	5960	5800	5879	6294	23933
	Percent (SE)	94.9 (0.29)	94.9 (0.28)	94.8 (0.28)	95.5 (0.28)	95.0

Table 3. Percent of Respondents Where the Actual Race/Ethnicity Outcome Matched the Substratum by number of Months Between C2SS Interview and NESARC Interview.

Race/Ethnicity Group		Number of Months Between C2SS Interview and NESARC Interview			Total
		6-8	9-10	11-17	
Hispanic	Number	2247	2586	2660	7493
	Percent (SE)	77.5 (0.89)	76.9 (0.80)	78.7 (0.77)	76.2
Black	Number	2352	2812	2421	7585
	Percent (SE)	90.4 (0.60)	89.6 (0.58)	86.9 (0.68)	89.1
Other	Number	6970	8964	7999	23933
	Percent (SE)	95.5 (0.24)	95.3 (0.22)	94.2 (0.25)	95.0

Table 4. Response Outcome by Race/Ethnicity Substratum.

				Other			
		Hispanic	Black	Valid race response	Vacant/ Missing	Group Quarters	Total
Completed Interviews	Number	10603	9130	25213	1881	280	47107
	Weighted[6] Percent (SE)	81.8 (0.40)	80.1 (0.40)	82.2 (0.22)	45.7 (0.78)	56.7 (2.40)	78.3 (0.19)
Eligible Non-respondents	Number	1203	1120	3268	410	51	6052
	Weighted Percent (SE)	9.8 (0.30)	10.2 (0.30)	10.7 (0.18)	10.0 (0.47)	10.3 (1.36)	10.5 (0.14)
Temporarily Ineligible	Number	901	959	1991	1596	86	5533
	Weighted Percent (SE)	7.7 (0.28)	9.0 (0.28)	6.5 (0.14)	38.7 (0.77)	17.4 (1.83)	10.0 (0.14)
Permanently Ineligible	Number	97	76	186	234	77	670
	Weighted Percent (SE)	0.8 (0.09)	0.7 (0.08)	0.6 (0.04)	5.7 (0.37)	15.6 (1.63)	1.3 (0.05)
Total	Number	12804	11285	30658	4121	494	59362
	Percent	100.0	100.0	100.0	100.0	100.0	100.0
	Response Rate (%)[7] (SE)	89.3 (0.33)	88.7 (0.33)	88.5 (0.20)	82.1 (0.81)	84.6 (1.98)	88.2 (0.16)

Table 5. Response Outcome by C2SS Mode of Data Collection

		Mail	CATI	CAPI	Total
Completed Interviews	Number	23159	9502	14166	46827
	Weighted Percent (SE)	84.8 (0.24)	82.7 (0.42)	68.5 (0.36)	78.5 (0.19)
Eligible Non-repondents	Number	2452	1141	2408	6001
	Weighted Percent (SE)	8.9 (0.19)	10.8 (0.34)	12.3 (0.25)	10.5 (0.14)
Temporarily Ineligible	Number	1640	685	3122	5447
	Weighted Percent (SE)	5.7 (0.16)	6.0 (0.26)	17.1 (0.29)	9.9 (0.14)
Permanently Ineligible	Number	179	54	360	593
	Weighted Percent (SE)	0.6 (0.05)	0.5 (0.07)	2.1 (0.11)	1.1 (0.05)
Total	Number	27430	11382	20056	58868
	Percent	100.0	100.0	100.0	100.0
	Response Rate (%) (SE)	90.5 (0.20)	88.5 (0.37)	84.8 (0.30)	88.2 (0.16)

[4] Group quarters cases were excluded from Tables 2,3, and 5.

[5] All standard errors in this paper were calculated using the SUDAAN software, which uses a Taylor series linearization method of variance estimation. All comparisons were tested at the 5% significance level.

[6] Base weights were used for all weighted calculations in this paper. Final weights were still unavailable at the time of publishing.

[7] The response rate is weighted and calculated as all completed interviews out of all eligible cases (completed interviews plus eligible nonrespondents).