COLLECTING HEALTH DATA ON MINORITY POPULATIONS IN A NATIONAL SURVEY

James T. Massey[1], David Judkins[2], and Joseph Waksberg[2]
[1]National Center for Health Statistics, [2]Westat, Inc.
James T. Massey, 6525 Belcrest Road, Hyattsville, MD 20782

Key Words: Oversampling minorities, alternative designs

INTRODUCTION

One of the problems often encountered in multipurpose national probability household surveys is the lack of precision for small domain statistics. In health surveys racial and ethnic comparisons are important for understanding health differentials within the population. Even in large surveys, however, the ability to produce reliable statistics for minorities is almost impossible. One of the largest ongoing national health surveys, the National Health Interview Survey (NHIS), has struggled with this problem for several decades. For its next redesign, scheduled for 1995-2004, one of the highest priorities was the improvement in the precision of minority statistics. A number of alternative designs were investigated for improving the precision of minority statistics. Many of these alternative designs are applicable to any multi-stage national probability survey.

Before discussing some of the alternative methods of improving the precision of minority statistics, it is important to study the distribution of the minority populations within the U.S. population. Tables 1 and 2 present the magnitude and concentrations of the population for the four largest racial and ethnic minorities in the U.S. The size of these groups vary from less than 1 percent for American-Natives to 12 percent for the Black population. Table 2 shows that the Black population is the most highly concentrated minority group followed by Hispanics, Asian-Pacific Islanders, and American-Natives. Both the size and distribution of the minority populations have major implications for oversampling these minority populations at a reasonable cost. For example, the fact that 60 percent of the total Black population live in Census defined blocks that contain 60 percent or more Blacks makes it straight forward to identify these blocks and oversample the Black population. On the other hand, the American-Native population is not so easily oversampled. While 30 percent of the American-Natives live on reservations or in highly concentrated areas of American-Natives, the remainder of these American-Native populations live in areas that contain 90 percent or more nonAmerican-Natives. This latter situation makes it almost impossible to effectively oversample the American-Native population in household probability samples. The patterns of residential concentrations by race and Hispanic origin are discussed by Judkins, Waksberg, and Massey (1992).

METHODS CONSIDERED FOR OVERSAMPLING MINORITIES

Simple Expansion

The least complicated approach to oversampling minorities is to simply increase the sample size. This option would require no changes in methodology. Unfortunately, it is an extremely expensive option. Research by NCHS and Westat concluded that for the NHIS, an effective sample size of approximately 1000 persons per age-sex subdomain would be required to produce most of the key NHIS statistics with a coefficient of variation (cv) of less than 30 percent. The precision requirements were also specified in terms of making comparisons between subdomains. To meet the precision requirement for elderly Hispanic males by this technique alone, the NHIS would require between a three-fold and four-fold increase in the NHIS sample size. Instead of interviewing 50,000 households per year, it would be necessary to interview 150,000 to 200,000 households per year. The numbers are even larger for elderly Asian-Americans and Native-Americans. Clearly, this is not a feasible technique.

One method of improving the precision of minority statistics through expansion is to combine several years of data. The NHIS design is ideally suited for combining data over time since each week's sample is a representative national sample. Two important assumptions are necessary, however, to combine multiple years of data. The survey must collect the data of interest in each of the combined years, and one must assume that the variable of interest does not change significantly over the combined years. The first assumption is rarely met in the NHIS since most of the topics of interest change each year.

Oversampling Classes of PSUs

It is well known that some non-self representing (NSR) primary sampling units (PSUs) have disproportionate numbers of Hispanics while others have disproportionate numbers of Blacks. This led to the idea that it might be advantageous to deliberately assign a larger measure of size to such PSUs so that more of them would appear in the sample. This would reduce between-PSU variance for minority statistics and make it easier to oversample minorities at the block level while maintaining reasonable PSU workloads. The PSU workload is the total number of households to be interviewed in a PSU over the course of a year. On the other hand, between-PSU variance is increased for others and for totals.

Since a very high percentage of Blacks and Hispanics live in large metropolitan areas that are not subject to between-PSU variance (their PSUs are SR), since between-PSU variance is already small for these domains, and since equal workloads are not an important consideration for NHIS with its traveling interviewers (equal monthly workloads are important but that can be controlled by the number of times that the PSU is visited over the course of the year), it was decided not to oversample PSUs with large minority populations. More details can be found in Fahimi and Judkins (1991).

Oversampling Classes of Blocks or Block Groups

Oversampling classes of blocks or block groups is the technique that was used in the 1985 NHIS redesign to oversample Blacks. Small areas that are known from the prior decennial census to be rich in minorities are deliberately oversampled. Unless combined with screening (see next section), this technique leads to variations in the probability of selection not just within the targeted domain but also within other domains. As a result, Whites who live in racially mixed neighborhoods were more likely to be selected for the NHIS than Whites living in predominantly White neighborhoods. This can lead to large design effects for statistics for Whites and the total population. This method was consequently also rejected.

Oversampling integrated and predominantly minority blocks is more effective for Blacks than for Hispanics since they are more segregated (Judkins, Massey and Waksberg; 1992). It is not very effective for Asian and Pacific Islanders since they tend to live in more integrated neighborhoods (ibid). Among Native-Americans, Eskimos, and Aleuts, there are two distinct patterns. Those that live on reservations are extremely segregated and thus easy to oversample. Those that live off reservations tend to be only slightly segregated (ibid).

Oversampling low median income blocks is not an effective procedure for oversampling the poor. Although there are obvious pockets of poverty, very significant numbers of the poor are dispersed throughout blocks and block groups.

Oversampling blocks is inefficient for oversampling different age groups such as teenagers or the elderly.

Screening with Subsampling

With screening, a larger sample is selected than is actually desired. A very brief (and hopefully) inexpensive questionnaire is then administered to the sample to determine whether any household members are in the domains to be oversampled. Those units (persons or housing units) that are not part of that domain are subsampled. It is possible to set up multiple domains with different subsampling rates. The subsampling can be done centrally after the completion of the entire screening operation, or it can be done by the interviewer during the household interview. Techniques have been developed that make the subsampling process very easy for the interviewer (Mohadjer and Waksberg, 1991). Interviewers don't need random numbers. Instead, they are given house-by-house pre-interview instructions about which domains can be interviewed at which households. These instructions are randomized centrally prior to screening to yield the desired sampling rates. Alternatively, with CAPI, the sampling can be done in other ways not transparent to the interviewer.

Experience at Westat has shown that it is possible to obtain basic information on sex, race, ethnicity, and age for the occupants of over 99 percent of housing units. A success rate this high requires persistent well-trained door-to-door interviewers and some recourse to proxy information through neighbors, but is definitely feasible.

Discussions with the Census Bureau resulted in a cost estimate of a single attempted screening equal to one third of a completed interview. Thus, for example, 51,000 extra attempted screeners should cost as much as 17,000 completed interviews.

Screening is most effective when combined with oversampling at the block level as is explored more in a later section. Screening can be used either at the person level or at the household level. Given the very low marginal cost in the NHIS of collecting data on extra members of a household, most of our attention was on screening whole households in or out.

Network Sampling (by nomination of relatives)

Prior applications of network sampling involved having an initial probability sample of persons (obtained by area or list sampling) report on the number of relatives they have outside the household and whether any of them have certain characteristics of interest such as diabetes (e.g., et al. (1978). As part of the NHIS redesign research, a variant of this method was studied for oversampling minority populations. It was believed that most persons would not be able to answer most of the questions in the NHIS accurately for relatives not living in the household. In fact, for many NHIS supplements, even a household respondent is considered unsatisfactory; interviewers are instructed to accept nonresponse instead of proxy response if self response is impossible. Thus, the scheme proposed and studied for NHIS was to have interviewers track the nominated relatives and administer the NHIS instruments directly with them. A possible refinement included the use of the telephone for the tracking and follow-up interviews.

This scheme is still under consideration for future redesigns but was dropped from active consideration for 1995 for several reasons. Most important of these reasons, was the newness of the scheme. There was insufficient time to pre-test it. Also, there is only limited data on the sizes of familial networks by race and ethnicity. The first step in carrying research forward on this idea would be to add a supplement to some major national survey that would fill in this information gap.

Dual-Frame Sampling (with Administrative Lists)

The only lists of persons that are maintained by the federal government that are reasonably complete and are available for sampling for health related research are the list of social security and railroad retirement act beneficiaries maintained by SSA and the list of medicare beneficiaries maintained by HCFA. Use of these lists falls under provisions of the Privacy Act of 1974 which permits transfer of the data (including names and addresses) within the Public Health Service for health related research. The SSA or HCFA lists can be used to select a supplemental sample of persons over 65 or persons receiving SSA benefits.

Commercial lists do exist for other segments of the population, but we have not studied their use.

Integration of samples from multiple frames into a single microdata file with a single weight requires, at a minimum, the ability to tell which units had dual chances of selection. This is enough to create composite weights. Furthermore, if it is possible to determine the sampling stratum for each unit in each of the frames, then the precision of the estimators can be improved. If the full joint probabilities of selection for all units across all frames can be determined, then even more precise estimators are possible.

BEST TECHNIQUES FOR NONELDERLY BLACKS AND HISPANICS

For the 1985-94 design, Blacks were oversampled by oversampling EDs and BGs with strong concentrations of Blacks. This procedure was less effective than it could have been since it was coupled to a policy of using total population as the measure of size for the selection of PSUs and keeping tight constrictions on total PSU workloads. This resulted in the situation where some rural southern PSUs with very strong concentrations of Blacks were not oversampled because it would have resulted in too large of a workload. The procedure could have been more powerful if the PSU measure of size had emphasized the presence of Blacks or if the constraint on workloads had been relaxed.

There is another inherent problem with the strategy, however. It results in an oversample of Whites, Asians, Native-Americans, and others who live in predominantly Black neighborhoods. Although a relatively small proportion of the nonBlack population lives in such neighborhoods, given the total size of the nonBlack population and the fact that nonBlacks in these neighborhoods aren't of greater policy interest than other nonBlacks, this oversampling results in a waste of effort.

Nonetheless, if the goal was just to oversample Blacks or just to oversample Hispanics, reasonable strategies could be built around oversampling blocks or block groups without screening. The problem comes from the objective to oversample both minorities. They tend not to live in the same areas. The blocks with strong concentrations of Blacks have only a small proportion of the Hispanic population and vice versa. Oversampling both Blacks and Hispanics requires a sharing of resources to oversample these minority populations and increases the proportion of Whites living in the minority areas (this leads to a greater variability in the sampling weights for Whites). Thus, simple oversampling without screening is not a viable method to meet the goals for the 1995 redesign. Table 3 provides a comparison of the effective sample sizes for oversampling with and without screening of households and clearly shows the advantage of oversampling with household screening.

Screening without oversampling blocks is probably the only technique that makes sense to oversample domains that are fairly evenly spread throughout the land such as teenagers or the elderly. However, for groups such as Blacks and Hispanics that are highly concentrated, screening by itself is not a good technique.

The problem is the expense. The Census Bureau estimated that each screening interview that did not result in a household of the desired type would cost as much as one third of the cost to interview the household with standard NHIS instruments. The final cost will be sensitive to the required response rate, the rules allowing the use of information from neighbors, and the complexity of the screening rules.

By oversampling at the block level and screening for race and ethnicity, significant simplifications and cost savings are achieved. Screening still has enormous benefits over oversampling blocks without screening due to the reasons cited above.

To set the details of the design, NCHS hopes to get a budget 50 percent larger than needed for a sample of 50,000 households in the current design. This would permit roughly equal effective sample sizes for Blacks and Hispanics, after allowing the sample size for Whites and other race/ethnic groups to fall in order to obtain large enough effective sample sizes for minorities. Large enough was defined to be about 8000 adults given the random selection of one adult per household. This rule was based upon the NCHS plan to administer most questionnaire supplements to just a single adult per household. Since the supplements are more burdensome than the core and since self-response is required from them, this subsampling does indeed save money.

Oversampling and screening in concentrated Black and Hispanic blocks permitted us to meet these specifications. Table 4 shows some aspects of the optimal mix of screening and oversampling for various cost assumptions and racial/ethnic groups. It also shows results separately for the observed prevalence and degree of clustering for each group in 1990 and for the projected prevalence and degree of clustering for each group in 2000. The latter is important given the deterioration of the benefits of stratification as the stratifying information become out-of-date.

There are five distinct panels within the table. The first panel shows how many times larger the screener sample should be than the interview sample. The second panel shows the ratio of the optimal sampling ratio in the densest stratum (with respect to population belonging to the group of interest) to the least dense stratum[1]. The third panel shows the size of the screener sample that would be required (in conjunction with the indicated stratification and optimal allocation) in order to achieve and interview sample with the same precision as a simple random sample of 1000 persons. The fourth panel shows the size of the required interview sample. These sizes are larger than 1000 due to the design effect induced by the differential sampling of the density strata. (Note that as a result of rounding some of the required interview sample sizes appear to be 1000. The exact numbers are all larger than 1000.) The fifth panel shows the benefit of differential sampling rates for the density strata. The basis of comparison is the hypothetical plan of simply screening a big enough sample of persons to find 1000 persons of the desired race/ethnicity.

To illustrate the interpretation of these panels, take for example the goal of oversampling Native-Americans. Based on 1990 data, if a full interview costs three times as much as a screener interview, then the screener sample should be 23.7 times larger than the intended interview sample. The screener sample should be selected at a rate 11.2 times higher in the densest Native-American stratum to the least dense stratum. It would be necessary to screen a sample of 61,000 persons and interview the resulting 2600 Native-Americans in order to match the precision of a simple random sample of 1000 Native-Americans (assuming hypothetically that such a sample were possible). Differential sampling rates lead to a 48 percent cost saving compared to straight screening of the general population.

All the numbers in Table 4 were derived using standard Neyman allocation assuming the cost ratios shown at the top and uniform population variances across the strata. The cost of one complete in a particular stratum is proportional to

$$\frac{\text{(Total persons in stratum)}}{k\text{(Targeted persons in stratum)}} + 1 - 1/k,$$

where k is the ratio of the cost of one full interview to one screener interview. (The constant of proportionality is not required to be known in order to carry out Neyman allocation.) To motivate this cost formula, consider an example. Suppose that a particular stratum, only one person in 100 was of the desired group and that the ratio of interview cost to screening cost is 20. Then, in order to get 1 interview, it is necessary (on average) to screen 100 persons, costing 5 "units." The interview itself costs 1 unit. A cost of 0.05 units is deducted to account for the fact that when a screened unit is carried forward into a full interview, the cost is not greater than a simple interview[2].

Reviewing the content of Table 4, several observations can be made. First (and somewhat obviously), as the cost of screening goes down, it makes sense to have a larger screener sample and depend less on oversampling by density stratum. The benefits of oversampling by density stratum also decline as the information on density becomes older. If full interviews are extremely expensive (because of physical examinations of subjects, laboratory work, multiple measurements of each case or other reasons), oversampling by density stratum is probably ill-advised except for the rarest racial/ethnic groups.

Second, because of the anticipated increase in the Hispanic population by the year 2000, smaller screener and interview samples are projected even though the stratification information will be older.

Third, Asian and Native-Americans are still so rare and so little segregated that enormous screener sample sizes are required for every interview. For most surveys, it will not be practical to try to provide high precision for these groups.

CHOOSING A TECHNIQUE FOR ELDERLY BLACKS AND HISPANICS

The number of elderly Black or Hispanic persons in this country is too small to produce sample sizes from the proposed design that meet NCHS requirements. Although there are projected deficiencies in the sample sizes for both sexes, the problems are especially acute for males. A budget increase much larger than 50 percent would be necessary to meet the goals through oversampling blocks and screening alone.

Since the Social Security Agency (SSA) maintains files with excellent coverage of the elderly population (at least for those 66 and older) we have been researching dual-frame sampling that would combine the traditional area/permit sample with a supplemental list sample. The task of oversampling elderly Blacks from SSA files is fairly straightforward since race is indicated for about 97 percent of the file. Unfortunately, SSA files do not have an indicator for Hispanic origin for persons that are currently elderly[3].

Using Surnames to Identify Hispanics

In order to use the SSA files for oversampling of elderly Hispanics despite the lack of an indicator for Hispanic origin, we considered the strategy of oversampling on the basis of surname, using the Hispanic surname file developed by J. Passel and D. Word at the Census Bureau (Passel and Word, 1980) for the purpose of classifying surnames by ethnic origin. The Passel-Word file contains 12,497 surnames that tend to belong to Hispanics.

Of course, the precision and cost of a dual-frame sample based upon surnames depends strongly on the sensitivity and specificity of the Passel-Word file. Every false positive costs money to interview or screen out and every false negative increases the portion of the area/permit sample that is not covered by the list, and has a high sampling weights, thereby increasing design effects due to unequal weights. A past study (Passel and Word, 1980) indicated that false positives (also called errors of commission) run at around 15 percent and that false negatives (also called errors of omission) run at around 20 percent. Given the time lapse since that original study, we thought it prudent to repeat the study. To that end, we undertook the matching of the surname of every member of the 1988 NHIS against the Passel-Word file. We then compared the self-reported ethnicity with the ethnicity that would be imputed on the basis of surname using this file. At the same time, we analyzed some of the characteristics of persons with nonconforming names, information useful for decisions about sample allocation.

The overall false positive rate is 12.6 percent while the overall false negative rate is 31.6 percent, as shown in Table 4A. The former rate is lower than we had expected based upon previously published research; while the later is higher than expected. Passel and Word originally reported error rates with the 1976 March Current Population Survey (CPS) of 15.0 percent false positive and 20.7 percent false negative. Part of the discrepancy involves persons who did not report their ethnicity, who had a hyphenated name or who refused to provide their name, all of which Passel and Word treated differently than we did. An inspection of nonconforming surnames of Hispanics indicates, however, that even allowing for these variations in matching procedure, there is a substantial and unexplained difference in the false negative rate. The reason for the poorer performance could either be lower quality recording, transcription, and keying of names on the 1988 NHIS than on the 1976 March CPS (interviewers, not respondents record name spellings) or real change in the population in the relationship between surname and self-reported ethnicity. Both rates are higher for the elderly: 16.6 percent false positives and 33.6 percent false negatives.

Differences are far more common among women than among men. Intermarriage evidently plays a strong role in both error rates. Socio-economic status also plays a strong role in both error rates. The general trend seems to be that higher socio-economic status means a weaker association between surname and self-reported ethnic origin. There are also differences by metropolitan status, part of the country and country of origin (for self or ancestors). See Judkins, Massey, and Smith (1992) for more detail.

Implications for Sample Design

Although the estimated sensitivity of the Passel-Word Hispanic surname list was not as good as we had hoped, the idea of using it to create a list sample of elderly Hispanics still has good potential. The technique is still far cheaper than area sampling with screening. It was assumed that the biases associated with the SSA sampling frame would have approximately the same magnitude as the biases for an area sampling frame. In this section, we explain these implications more fully. Table 5 contrasts some of the numbers that appear in the following text.

Under current plans for 1995 and beyond, the area/permit NHIS sample will yield nominal elderly Hispanic sample sizes of about 500 males and 700 females. After accounting for the design effect due to disproportionate sampling of heavily Hispanic blocks, the effective sample sizes (compared to a similarly clustered sample with equal probabilities) will only be about 350 and 500. By adding 1000 males with Hispanic surnames from SSA lists and another 1000 females in the same manner, we can boost the nominal sample sizes to around 1240 and 1400 and the effective sample sizes to around 770 and 920. (The effective sample sizes don't increase as much as the nominal sample sizes because of the design effect due to the large weights that Hispanics without Hispanic surnames will bear.) To get a comparable boost from the area/permit sample alone would require the screening of an additional 100,000 households!

As another contrast, suppose that one used SSA lists as a supplemental list sampling frame without paying attention to surname. In that case, more than 20,000 persons would have to be located and screened on ethnicity in order to get comparable boosts for elderly Hispanics. Thus, although screening a list of elderly persons is far more efficient than screening a sample of households, even if those households were heavily skewed toward heavily Hispanic blocks, it is nowhere near as efficient as using a list in combination with the surname list.

Nonetheless, we had hoped to boost the effective sample sizes even more sharply for elderly Hispanics, to as high as 1000 males and 1000 females. Here it turns out that the false negative rate is too high to make that economical. About 5000 list persons with Hispanic surnames would have to be added to the sample.

We are thus extremely interested in determining whether there are methods of decreasing the false negative rate, even if it means some increase in the false positive rate, as it likely would. If, for example, the 1976 findings of Passel and Word still held, the supplemental sample size required for the desired effective sample sizes by sex would be just 2400 instead of 5000. We suspect that adjustments in interviewer training and in keypunching could reduce the frequency of false negatives. (Even though we would be sampling from an SSA list that would probably have higher quality name spelling, we would still need to classify everyone in the area/permit sample in order to work out appropriate sampling weights for the dual-frame estimator.) It is possible, however, that there has been a sea-change in the relationship between Hispanic origin and surname and that improvements will be difficult. David Word (in a personal communication) doubts that such a change has occurred. He is conducting research similar to ours on the much larger sample in the Census Post Enumeration Survey. It will be very interesting to compare results. Table 6 presents a summary of the potential gains in precision of using a supplemental SSA list sample in the NHIS. The gains in precision depend upon the coverage of the domain of interest by the list sampling frame.

PROSPECTS FOR STATISTICS ON DETAILED HISPANIC SUBDOMAINS AND OTHER MINORITIES

Hispanic Subgroups

Despite the progress that the proposed design would make in improving the precision of NHIS statistics about all Hispanics, the diversity of the Hispanic subgroups is so great, that it is necessary to break the statistics further for truly helpful analyses. There is interest in separate statistics on Cuban Americans, Puerto-Rican Americans, Mexican Americans (separately by whether their ancestors lived in the U.S. prior to 1900 or not), Central America Americans, and other Hispanic Americans.

Although this may be true, there are definite limits imposed by a fixed budget. Given the overwhelming importance of age and sex for questions related to health, it is clear that valid contrast between detailed Hispanic subgroups would have to be supported by adequate sample sizes for each age and sex subgroup within each Hispanic subgroup. Assuming 8000 per detailed subgroup, five detailed subgroups would require 40,000 persons, about one-third of the total NHIS sample. An even larger proportion of the NHIS budget would be required.

Our research showed that it is not possible to obtain design-based statistics of the desired reliability about detailed Hispanic subgroups with the current budget. Cuban-Americans, in particular, are too rare and dispersed. Even a doubling or trebling of the budget would allow only modest precision.

We also investigated the possibility of breaking Hispanic statistics geographically instead of ethnically. This approach has been used in the past for Hispanic HANES (Health and Nutrition Examination Survey). The regions that we investigated were 1) the New York CMSA, 2) Florida, and 3) the five southwestern states of California, Arizona, New Mexico, Texas and Colorado. This investigation also showed the need for budget increases far larger than can be expected.

On a more positive note, there are several ways to produce some reliable statistics for the major Hispanic subgroups. By combining multiple years of data, detailed statistics by age and sex are possible for most of the major Hispanic subgroups. Some statistics for the Hispanic subgroups have been produced from the current NHIS design (see Moss (1984)). With the proposed new design, even more statistics could be produced. The second way to provide statistics for the Hispanic subgroups is to make the data requirement less stringent. Under the new design it will be possible to produce marginal statistics for the Hispanic subgroups by collapsing age groups or combining males and females where appropriate.

Given the limitation of design-based techniques, we also looked at model-based techniques. In this research, we considered the possibility of oversampling only Hispanics in high Hispanic blocks and then extrapolating the results to all Hispanics, using models. A key factor here is the relationship between health and the degree of residential clustering. If the clustering is unrelated to health or if the relationship can be explained by socio-economic variables that are available from other data sources, then the model-based approach would be unbiased and much more economical than design-based estimators. Even if neither of these conditions apply (and it seems natural to believe that there is some relationship that can't be entirely explained away in terms of other variables) and thus the model-based estimates would be biased, it is possible that model-based estimators would still have acceptable mean square error.

To study the question, we analyzed health data for Hispanics in the 1988 NHIS. There were too few to study the detailed groups separately. We contrasted the health characteristics of Hispanics across the five minority density strata that we had created for Hispanics. (The strata were defined as of 1980.) Simple comparisons showed significant differences. We then created logistic models for three binary health measures. The inclusion of socio-economic variables reduced the nominal relationship between residential clustering and health but did not rule out the existence of important residual relationships. The sample sizes weren't really large enough to make very firm projections of the mean-square error for model-based estimators, but the results were not encouraging.

Asians and Pacific Islanders

The main problem in sampling Asian and Pacific Islanders is that despite the areas like the Chinatowns of New York and San Francisco, most of them live in integrated neighborhoods where it is expensive to find them. A further problem is the diversity of the populations covered by the OMB classification of "Asian and Pacific Islander." Just as there is a need for Hispanic statistics by subgroup there is a need for Asian-american statistics by subgroups such as Chinese, Japanese, Polynesians, Indians, Malays, Vietnamese, Hmong, Laotians, and so on. Median income is now higher for Asians and Pacific Islanders than for Whites. It is not clear that there are larger health-related differences between Asians and Pacific Islanders on the one hand, and Whites on the other, than there are among the various Asian subgroups.

Based upon statistics from the 1990 Census, we estimate that a modest oversample of Asians and Pacific Islanders could be achieved by screening an additional 15,000 households (over the 99,000 already planned) and by interviewing an additional 2600 households. The resulting oversample would yield an effective sample size of around 8000 persons. Effective sample size here means the normal sample size divided by the design effect due to unequal probabilities but not due to clustering. This size of 8000 persons could only be achieved by interviewing all members of Asian households. (Asian households are simply too rare to allow subsampling within households to reduce household burden.) Also, the distribution across ages would not be optimal. Narrow age bands of particular analytic interest such as persons 18 to 24 and children under age 6 would have smaller sample sizes than desired.

Network sampling may have some promise for Asians. To develop the idea further would require a supplement to some large national survey (not necessarily NHIS) where Asians are asked for the names, addresses, and phone numbers of parents and siblings. Such a supplement would allow more accurate projection of the multiplicities for various network rules. A field test would, of course, have to precede such a supplement to fine-tune the questions and determine whether the supplement could have any negative effects on overall response propensity.

At least one or two companies are experimenting with lists of typical Asian surnames as a means to identity Asian and Pacific Islanders. If such an approach yields a false negative rate under 40 percent, then at least modest oversampling would be possible.

Multi-year aggregation is also a technique to be considered for Asians and Pacific Islanders. In conclusion, however, there was the feeling that the funds necessary would amount to such a large part of the total NHIS budget that it was not practical at this time.

American Indians, Eskimos and Aleuts

The relatively small percentage of Native-Americans makes it quiet expensive to locate a sufficient sample size for useful statistics. About half of them live on reservations where they are easy to find, but the other half live in mostly well integrated neighborhoods where finding them requires extensive screening. If the goal was just to provide reliable statistics about Native-Americans living on reservations, this goal could probably be achieved with a modest budget increase (perhaps less than 10 percent).

Otherwise, we estimate that a screening sample of an additional 128,000 households (on top of the planned 99,000) and an additional 5860 household interviews would be required just to give an effective sample size of 8000 Native-Americans, Eskimos and Aleuts, even if all household members are interviewed.

DISCUSSION OF DESIGN OPTIONS FOR THE NHIS

After completing the research on the different design options for oversampling Blacks and Hispanics, the parameters for the final design had to be selected. It was clear that the best design alternative for improving the specified level of precision of statistics for Blacks and Hispanics assuming a 50 percent increase in the NHIS data collection budget would involve oversampling high density minority blocks with household screening. Assuming a 50 percent increase in the NHIS budget for the next decade is questionable at best. In order to make a final design decision and have some back-up design options, a number of different design alternatives were evaluated. Six of the design options are shown in Table 7. The relative effective sample design for Blacks, Hispanics, and others are shown for several design options and budget assumptions. The fully funded NHIS sample would double the precision for Black statistics and triple one precision for Hispanics. The loss in precision for "others" would only decrease by 23 percent and still be much greater than the precision for the two minority populations. With no increase in the NHIS budget the improvement in the precision of the minority statistics can't be realized, even if the precision for Whites is decreased by 80 percent. This implies that under the current NHIS budget, precision of minority statistics can't be significantly improved without eliminating many other important cross-classifications of the NHIS data. It does not appear that the precision of Hispanic statistics could be increased to the level of precision for Blacks by reducing the precision of statistics for "others" and the "total" population by approximately 50 percent. This option (2) was selected as the back-up option for the NHIS under the assumption of no increase in budget.

Table 1. **U.S. population by race and ethnic origin in 1980 and 1990**

Race or Ethnic Origin	1980		1990	
	Number (in thousands)	Percent of total	Number (in thousands)	Percent of total
Black	26,495	11.7	29,986	12.1
Asian or Pacific Islander	3,500	1.5	7,274	2.9
Native-American, Eskimo, or Aleut	1,420	0.6	1,959	0.8
Other Race	195,130	86.1	209,491	82.4
Total	**226,546**	**100.0**	**248,710**	**100.0**
Hispanic	14,609	6.4	22,354	9.0

Table 2. **Concentration of minority populations in strata defined by percent minority in block or block group in 1990**

Stratum (Designated minority as % of Block or BG)	Percent of Designated Minority in Each Stratum							
	Hispanic		Black[1]		Asian or Pacific Islander[1]		Native-American, Eskimo, Aleut[1]	
	Block	BG	Block	BG	Block	BG	Block	BG
0-10	14.8	19.2	8.3	11.7	37.0	47.8	46.8	57.8
10-30	22.1	22.8	13.6	16.5	32.1	27.8	15.9	12.4
30-60	23.3	24.1	16.0	20.1	18.0	14.6	7.7	6.0
60-100	39.8	33.9	62.2	51.6	13.0	9.8	29.6	23.8

[1]Non-Hispanic

Table 3. **Effective sampling sizes for several sampling options at constant cost assuming a cost ratio of 3:1 for interviewing versus screening**

	Self-weighting 50,000 HHD completes		Oversampling w/o Screening		Oversampling with Household Screening for <65	
Blacks						
	Males	**Females**	**Males**	**Females**	**Males**	**Females**
Under 5	689	660	857	821	1360	1303
5-17	1963	1873	2443	2331	3874	3697
18-24	728	852	906	1060	1437	1682
25-44	1971	2500	2453	3111	3890	4934
45-64	1264	1632	1573	2031	2495	3221
65 plus	584	864	727	1075	1153	1705
Hispanics						
	Males	**Females**	**Males**	**Females**	**Males**	**Females**
Under 5	704	674	933	893	2209	2115
5-17	1691	1615	2241	2140	5306	5068
18-24	641	620	849	822	2011	1946
25-44	1790	1747	2372	2315	5617	5482
45-64	804	911	1065	1207	2523	2859
65 plus	290	406	384	538	910	1274
Others						
	Males	**Females**	**Males**	**Females**	**Males**	**Females**
Under 5	2820	2673	2143	2031	2054	1947
5-17	8520	8080	6474	6140	6207	5887
18-24	4098	4135	3114	3142	2986	3013
25-44	14144	14346	10748	10901	10305	10452
45-64	11621	12065	8830	9168	8466	8790
65 plus	5735	8290	4358	6299	4178	6040

Table 4. Aspects of the optimal mix of screening and oversampling for various cost assumptions and racial/ethnic subdomains

	Ratio of Interview Cost to Screening Cost									
	1:1	1.5:1	2:1	3:1	4:1	5:1	10:1	20:1	40:1	60:1
Based on 1990 Distributions	Optimal Ratios of Screening Sample to Interview Sample									
Blacks	2.6	2.8	3.0	3.3	3.5	3.7	4.3	5.1	5.9	6.4
Hispanics	3.7	4.0	4.2	4.5	4.8	5.0	5.7	6.6	7.6	8.2
Asians and Pacific Islanders	10.1	10.8	11.3	12.1	12.8	13.3	15.4	18.1	21.2	23.2
Native-Americans, Eskimos and Aleuts	15.8	18.5	20.5	23.7	26.1	28.2	35.8	45.4	57.2	65.2
Based on Projections to 2000										
Blacks	3.1	3.4	3.6	4.0	4.2	4.5	5.2	5.9	6.6	6.9
Hispanics	3.9	4.2	4.4	4.7	4.9	5.1	5.8	6.5	7.1	7.4
Based on 1990 Distributions	Optimal Oversampling Rates for Densest Stratum									
Blacks	8.2	6.8	6.0	5.0	4.4	3.9	2.9	2.2	1.7	1.5
Hispanics	9.7	8.2	7.3	6.1	5.3	4.8	3.5	2.6	2.0	1.7
Asians and Pacific Islanders	10.7	9.1	8.1	6.8	6.0	5.4	3.9	2.9	2.2	1.9
Native-Americans, Eskimos and Aleuts	20.4	16.2	13.9	11.2	9.7	8.6	6.1	4.4	3.2	2.6
Based on Projections to 2000										
Blacks	5.5	4.6	4.0	3.4	3.0	2.7	2.0	1.6	1.3	1.2
Hispanics	5.6	4.7	4.1	3.4	3.0	2.8	2.1	1.6	1.4	1.3
Based on 1990 Distributions	Screener Sample Sizes for Precision Equivalent to Simple Sample of 1000									
Blacks	3,900	3,900	4,000	4,100	4,200	4,300	4,700	5,300	6,000	6,400
Hispanics	5,500	5,500	5,600	5,700	5,800	5,900	6,300	7,000	7,800	8,300
Asians and Pacific Islanders	18,000	18,000	18,000	18,000	19,000	19,000	20,000	21,000	23,000	25,000
Native-Americans, Eskimos and Aleuts	61,000	61,000	61,000	61,000	62,000	63,000	65,000	69,000	75,000	79,000
Based on Projections to 2000										
Blacks	4,700	4,700	4,800	4,900	5,000	5,100	5,600	6,100	6,700	7,000
Hispanics	5,500	5,500	5,500	5,600	5,700	5,800	6,200	6,700	7,200	7,500
Based on 1990 Distributions	Interviewed Sample Sizes for Precision Equivalent to Simple Random Sample of 1000									
Blacks	1,500	1,400	1,300	1,200	1,200	1,200	1,100	1,000	1,000	1,000
Hispanics	1,500	1,400	1,300	1,300	1,200	1,200	1,100	1,100	1,000	1,000
Asians and Pacific Islanders	1,800	1,700	1,600	1,500	1,500	1.400	1,300	1,200	1,100	1,100
Native-Americans, Eskimos and Aleuts	3,800	3,300	3,000	2,600	2,400	2,200	1,800	1,500	1,300	1,200
Based on Projections to 2000										
Blacks	1,500	1,400	1,300	1,200	1,200	1,200	1,100	1,000	1,000	1,000
Hispanics	1,400	1,300	1,300	1,200	1,200	1,100	1,100	1,000	1,000	1,000
Based on 1990 Distribution	Percent Reduction in Cost Compared to Screening without Oversampling									
Blacks	53	48	43	36	31	27	16	7.6	3.1	1.7
Hispanics	51	47	43	38	33	30	19	10.2	4.6	2.7
Asians and Pacific Islanders	47	45	43	41	38	36	28	18.8	10.4	6.6
Native-Americans, Eskimos and Aleuts	52	51	50	48	47	46	40	33.1	24.3	19.0
Based on Projections to 2000										
Blacks	41	35	32	26	21	18	10	4.1	1.5	0.8
Hispanics	35	31	27	22	19	16	9	3.7	1.2	0.5

Table 5. Sensitivity and specificity of Hispanic surname as a surrogate for self-reported Hispanic origin by sex, age, and marital status

Characteristic	Sensitivity of Hispanic Surname		Specificity of Hispanic Surname	
	Estimated persons of Hispanic Origin	False negative rate	Estimated persons with Hispanic Surnames	False positive rate
	(in thousands)		(in thousands)	
Total	19,393	31.6	15,195	12.6
Male	9,452	27.7	7,597	10.1
Female	9,941	35.2	7,598	15.2
Age				
0-4	2,147	32.1	1,651	11.8
5-17	5,132	31.4	3,832	8.2
18-24	2,600	31.7	2,019	12.0
25-44	6,180	31.7	4,990	15.4
45-64	2,464	30.2	2,010	14.4
65+	871	33.6	694	16.6
Marital Status				
Under 14 yrs.	5,793	31.6	4,364	9.1
Married, spouse in household	7,302	31.1	5,973	15.8
Married, spouse not in household	280	22.7	242	10.5
Widowed	467	38.2	349	17.3
Divorced	773	38.8	68	16.6
Separated	414	31.6	332	14.8
Never Married	4,282	30.9	3,271	9.5
Unknown	83	30.8	6	40.2

Source: 1988 National Health Interview Survey (NHIS). 9600 NHIS sample persons reported Hispanic origin and 7500 NHIS sample persons gave a surname listed on the Passel-Word Hispanic surname file. Estimates shown in table are weighted to U.S. population.

Table 6. Required sample sizes by sampling method and precision target for Hispanics 65+ years, 1995 National Health Interview Survey

Method	Sex	Effective Sample Size	Supplemental Screener Interviews[1]
SSA List with match to Passel-Word Hispanic Surname File under observed error rates	M	770	1,000
	F	920	1,000
			2,000
Area Sample with household screening	M	770	
	F	920	
			100,000[2]
SSA List without match to Passel-Word Hispanic Surname File	M	770	9,750
	F	920	10,350
			20,100

[1]The optimal number of screener interviews depends on the cost of screening. For this work, it was assumed that a full interview costs three times as much as a screener interview.

[2]Number of additional households requiring a screener interview to identify Hispanic households. The expected number of households to be selected for the 1995 NHIS is 99,000.

Table 7. **Percent reduction in coefficient of variation (cv) for the National Health Interview Survey Disability Survey by size of supplemental SSA sample and the proportion of the subdomain of interest receiving SSA benefits**

Size of SSA Supplemental Sample	Proportion of Subdomain Receiving SSA Benefits					
	0.1	0.25	0.5	0.75	0.9	1
			(percent reduction in cv)			
2000	14	16	21	26	30	32
4000	12	17	25	34	40	45
8000	8	14	26	40	50	58
16000	-6	3	18	38	54	69
32000	-29	-18	2	29	52	77

Note: A fixed cost was assumed for the NHIS. Thus as the size of the SSA supplemental sample increases, the size of the area sample was decreased proportionately.

Table 8. **Effective adult sample sizes by design option and domain -- relative to self-weighting design in 1988 -- assuming one sample adult per household (design effects include effect of oversampling of high black EDs and within household subsampling, but not for clustering which is the same in all options)**

	Self-Weighting Design in 1988	Options for the Year 2000					
		1) 50% Funding Increase	2) Level Funding with Oversampling and Screening	3) 15% Funding Increase to Re-use Prior Year Minorities	4) 15% Funding Increase with Full Oversampling and Screening	5) Level Funding with Uniform Precision for Blacks, Hispanics, Others	6) Level Funding with Half Precision for Minorities vs. Others
Blacks	4,000	+113%	+35%	54%	113%	+72%	+53%
Hispanics	2,300	+282%	+142%	+105%	282%	+209%	+176%
Others	32,400	-23%	-51%	-1%	-94%	-78%	-61%
Totals	39,200	-17%	-48%	+5%	-95%	-80%	-65%

[1]For Hispanics, for Asians and Pacific Islanders and for Native-Americans, Eskimos and Aleuts, the least dense stratum was defined to be under 5 percent rather than the under 10 percent used for Blacks.

[2]This assumption would not be true if a separate contact was required to conduct the full interview. Most survey administrators avoid central sampling since the required second contact associated with central sampling gives the sample person extra opportunity to be a nonrespondent.

[3]SSA started to collect ethnicity in the 1980's, clearly too late to be of use in classifying the current elderly.