INTEGRATION OF THE SAMPLE DESIGN FOR THE NATIONAL SURVEY OF FAMILY GROWTH WITH THE NATIONAL HEALTH INTERVIEW SURVEY

Joseph Waksberg, Westat, Inc.

1. Survey Requirements and Implications

1.1 Outline of Research Carried Out

Dr. Sirken has given the background and described the purposes of the research carried out on this project, and it is not necessary to repeat it in detail. In brief, the research described in this paper was undertaken to develop alternate methods of selecting a sample of eligible women for the fourth cycle of the National Survey of Family Growth (NSFG) from the National Health Interview Survey (NHIS), consider survey methods that are possible with the samples, and analyze the characteristics of these methods. The ultimate goal, of course, was to provide NCHS with sufficient information on the alternatives to permit the agency to choose the one that appeared to be the best from an overall point of view.

Let me start off by outlining the survey requirements and their implications for sample design, and the steps necessary to examine sample designs that would achieve these requirements efficiently. I will then discuss these steps in greater detail. The main features of the project were as follows:

- The first step was to establish the survey requirements in terms of variance specifications. In practice, this amounted to examining the sample sizes considered for the next cycle (Cycle IV) of the National Survey of Family Growth (NSFG) with design effects that would apply if the sample design were used as in the two earlier cycles of NSFG.
- Secondly, other survey requirements and constraints were then listed.
- A set of alternative sample designs was then established which *a priori* seemed capable of efficiently meeting the survey requirements.
- Design effects were then estimated for each alternative and effective sample sizes calculated for each; that is, the sample sizes that would produce approximately identical sampling variances.
- Other major features of each sample design were then estimated, including the amount of screening that would be required, approximate total cost (excluding fixed costs of planning, analysis, etc.), the length of time it would take for HIS to generate the necessary samples, and expected effect of alternative methods on response rates.
- These features were then displayed in a form that simplified comparisons, to assist NCHS in choosing the one that seemed best on the basis of all features of the survey.

1.2 Sample Size and Sampling Rate Requirements with Use of Earlier Sample Design

The variance specifications for Cycle IV of the NSFG were that the sample design should produce the same sampling variances as would have resulted from a sample of 14,000 women selected in accordance with the Cycle III sample design. This equality of variances should also hold for four classes of women: Black ever-married women, Black never-married women, White and other race ever-married, and White and Other Race never-married. It should be noted that the survey is restricted to women in the main child-bearing ages, 15-44 years of age.

The sample sizes (with the Cycle III design) are shown in Table 1 together with the estimated population totals and the sampling rates that would be necessary with an equal probability sample. The last two columns of Table 1 oversimplify the situation. Other constraints on survey methods introduce even greater variability in the sampling rates than is shown in the last column. However, even ignoring the complications, the last column indicates the nature of the problem -- the fact that the sample size is increased by a factor of 4.55 to locate the sample of Black ever-married women, as compared to ever-married women who are White or of Other races. In practice, this means that with an equal probability sample, and ignoring other sample requirements, and allowances needed for nonresponses, it is necessary to screen about 17,500 households for the White ever-married sample, but about 79,500 for the Black sample. The additional complications increase these numbers even more. This is the requirement that introduced complications into the sample designs for the earlier cycles of NSFG and at least partially motivated the efforts to integrate the sample with that of the NHIS.

1.3 Other Requirements for NSFG, Cycle IV

Several other requirements for the NSFG affected the sample design. The main ones were:

- The total sample for Cycle IV is to come from NHIS. A dual-frame approach is excluded from consideration.

- In households containing two or more eligible women, no more than one should be included in the sample.

- Face-to-face interviews are to be carried out in personal visits. In effect this implies a multi-stage design with large PSU's as the first stage.

- The sample should be available in time to complete the interviewing in about 1986 or 1987.

- The response rates should not be appreciably worse than achieved in Cycle III. Cycle III had an overall response rate of 79 percent, composed of a 95 percent rate in screening and 84 percent in interviewing.

2. Alternative Designs Considered

2.1 NHIS Sample

A description of the main features of the NHIS sample will be useful in seeing how it can be adapted to NSFG. The NHIS is a survey of a sample of households in the United States. The annual sample is divided into 52 subsamples, with interviewing carried out weekly for the subsamples.

The NHIS is currently being redesigned and starting with January 1984, the annual sample will comprise 50,000 households, with Black households oversampled by 20 percent. This will bring the percent of Black households in the sample to about 14 percent of all households, or 7,000 of the 50,000. About 10 percent of the NHIS interviews will be carried out by telephone, and these will not be available for NSFG. The available annual sample thus will consist of 45,000 households, of which 6,300 will be Black and 38,700 will be White and Other races.

The NHIS sample will be a multi-stage design, the first stage consisting of 200 PSU's. The second stage of area segments will contain an average of eight housing units. However, this segment size applies only to the annual survey. The segments in adjoining years are usually next to each other. Consequently, if two years are combined, the effective segment size is 16 households. The third stage consists of households. Information is obtained about all persons in the sample households.

2.2 Alternatives Considered for NSFG Cycle IV

The alternatives considered comprised combinations of three basic variables:

1. A 200- vs. 100-PSU sample design. These two appeared to be the practical upper and lower limits possible. It is, of course, possible to develop samples within this range, but it seemed that an analysis of 100 and 200 PSU's would provide the data needed for such a decision.

2. Whether the sample frame should consist of eligible women in the NHIS with movers tracked and interviewed at their new residences, or whether it should consist of sample addresses. In the latter case, it would include addresses both with and without eligible women, although the former would be sampled at a much higher rate. New construction, vacant housing, and nonresponses in NHIS would be treated as part of the noneligible addresses. Since this would be a sample of addresses, no tracking would be necessary. Although this procedure would have higher design effects than using a sample of eligible women and thus require a larger sample for the same precision, the fact that it would not require tracking and following movers made it appear to be a competitive alternative.

3. The third variable was not one of sample design but of field interviewing procedures. Two possible methods were considered: The first is the accumulation of sample cases in NHIS until the desired sample size is attained before starting the field operations. The length of the interview period would then be the same as in earlier cycles of NSFG, about four months. The second method is continuous interviewing, in effect having the NSFG interview follow closely behind the NHIS interview. With this alternative, the interview period would approximately cover the length of time necessary to accumulate the sample from NHIS.

Eight possible basic designs exist consisting of all combinations of these three variables. Furthermore, these eight can be considerably expanded. The second variable provides for sampling housing units at different rates, depending on whether or not they contained eligible women in NHIS. Various combinations of rates are possible, leading to different alternatives. On the other hand, early in the course of the analysis it became clear that some combinations were unnecessary or impractical.

The alternatives thus narrowed down to seven. The Cycle III model, consisting of the design used in Cycles II and III, was an eighth included in the analysis, so that the alternatives using NHIS could be compared to each other and also to the design used earlier. The eight designs are shown in Table 2.

3. Design Effects for Alternative Designs

3.1 Components of Design Effects

The first part of the research was to ascertain the sample sizes needed for each of the alternatives. The various sample designs have different design effects leading to the need for different sample sizes for the same precision. The design effects arise from a number of features of the sample design -- principally, variability in sampling rates among certain segments of the population; subsampling in multi-eligible households; and the use of multi-stage sample designs involving between-PSU effects, between-segment effects, and variability in segment size.

Components of variance contributed by the first two features, variability in sampling rates and subsampling in multi-eligible households, were estimated in two steps. The first step was to develop the proportions of the population that would be sampled at each rate used in the survey, in some cases using data from Cycle II and in other cases, using various Census sources. The second step was to use these estimates to calculate increases in design effects. Kish has shown that under certain conditions

$$(\Sigma P_i k_i)(\Sigma P_i/k_i)$$

is a good approximation to the relative increase in variance arising from sampling at different rates, where P_i is the proportion of the population being studied sampled at a specific rate, and k_i is proportionate to the sampling rate (Kish, 1965). This approximation is quite good under a fairly broad range of conditions and was used to estimate the design effects arising from variability in sampling rates.

The components of design effects arising from multi-stage sampling were estimated through use of BRR variance estimation from Cycle II data, in which separate calculations were made of between-PSU variances, between-segment variances, and within-segment variances. Variances were calculated for 26 separate statistics, and average design effects were used. For each sample design (including the Cycle II and Cycle III model), the components of variance were combined to produce an estimate of the total design effect. Sample sizes necessary to produce a fixed level of precision were then calculated.

One factor causing variation in sampling rates was the constraint that only one eligible woman per household could be interviewed. This implied that women in households with two eligible women had only half the probability of selection of those in households with one eligible woman, etc. This constraint affected all sample designs about equally.

This is the main reason for deviation from an equal probability sample for designs 1 to 4, that is those based on a sample of persons in NHIS. The other designs have other features which turn out to have much greater effects than those caused by the one-person-per-household condition.

With the Cycle II and III design, most of the variability is caused by the need to reduce the extensive screening required for the Black sample. The sample designs for Cycles II and III used basically two rates for Black women. Two strata were established: one consisting of those Census block groups or ED's which were reported to have a fairly high proportion of Blacks in the most recent Census, the other consisting of the rest of the country. For Black women, the first stratum was sampled at a much higher rate than the second. (White women in the predominantly Black stratum were subsampled so that the disparity in their sampling rates was not as great.)

Designs 5 to 7 have a similar feature in that housing units that have eligible women at the time of NHIS are to be sampled at a rate 2 to 4 times the rate for other housing units. As a result, mobile women are undersampled relative to those tending to remain at the same address.

It is interesting to examine the proportion of the sample that would be included at different rates for the various designs. Table 3 shows the sampling rates used in Cycle II, rates that would be applied with the same design in Cycle IV, and the proportion of eligible women sampled at each rate. Table 4 has the design effects resulting from this disparity in sampling rates. The large design effects for statistics about Black women come primarily from the substantial differences in sampling rates and the fact that 28 percent of Black women are expected to live outside the predominantly Black stratum. The 28 percent is primarily due to the fact that the Cycle IV will be carried out in 1986 or 1987, and the 1980 Census data will be considerably outdated by then.

The effect of subsampling multi-eligible women households is not as great, but still far from trival. Table 5 shows the distribution by household size. Blacks and Whites are affected about equally. A much higher percent of never-married women would have to be subsampled. To reduce the design effects for never-marrieds, we have proposed to select them at twice the rate of ever-marrieds, when the two types occur in the same household.

With the use of the NHIS as a sampling frame, there is no need to create the four strata used in Cycles II and III. The only reason for differential sampling rates for designs 1 to 4, that is ones that involve following up persons who move, is subsampling in households with two or more eligible women. However, designs 5-7, which include all housing units with an eligible woman, and a subsample of the rest, will have strata with different rates. For these alternatives, the additional component of variance will depend on the subsampling rate and the number of years NHIS that are used. The number of years affects the design effects because the greater the time interval between NHIS and NSFG, the higher the proportion of eligible women that will have moved out of the stratum of addresses taken with certainty and into the one subsampled.

Design effects for these designs were estimated with the same Kish formula used for other causes of variability in sampling rates. The values of P_i, the proportions of the population sampled at different rates, came from Census data on population mobility. Table 6 contains the mobility rates used.

The design effects caused by multi-stage sampling were approximated by first estimating the intraclass correlations arising from sampling PSU's, and segments, and multiplying them by the average PSU and segment size for the alternatives. To calculate the intraclass correlations, a BRR program was used to estimate the total variances, within PSU variances, and within-segment variances from a microdata file of Cycle III. Differences between pairs then provided data on the increases in variance due to the two-stages of selection. Using the average size segment and PSU in Cycle III, these were converted to intraclass correlations.

3.2 Design Effects and Implications for Sample Size

It is interesting to see which components are the major contributors to design effects. Table 7 shows the individual components for the Cycle III design applied to Cycle IV sample sizes. For Black women, the major cause of the fairly high design effects is variability in sampling rates, with the variability among strata dominating. For White women, the design effects are about equally split between multi-stage sampling and differential sampling rates, with subsampling multi-eligible households the primary contributor to the latter.

Table 8 shows the design effects for the various alternatives. For the sample of housing units, r = 1/2, 1/3, 1/4 refer to the subsampling rates used for housing units without eligibles in the NHIS. It can be noted that the design effects are slightly affected by the number of years of NHIS that are needed to generate the required sample sizes. The reason is that combining several years of NHIS increases the segment sizes.

Design effects for most of the alternatives are considerably below those for the Cycle III design, particularly for Black women. This, of course, gets reflected in lower sample sizes needed for equal precision. These sample sizes are shown in Table 9. As can be seen, the use of NHIS permits considerable reduction in sample sizes.

4. Other Factors Affecting Decisions on Most Efficient Alternative

4.1 Number of Years NHIS Needed

The NHIS can only supply a limited sample size each year. A comparison of the sample size needed for NSFG and the number available annually from NHIS indicates how long it will take NHIS to generate the sample sizes required. The number of years is shown in Table 10.

The period of time is important for several reasons. First of all, the earliest NHIS can start accumulating a sample is January 1985. Since the interviewing is to be conducted in 1986 or 1987, three years of NHIS will not be available. Sampling of housing units with 100 PSU's is thus not feasible. In fact, there is some uncertainty about the use of 100 PSU's with a sample of persons. Secondly, aggregating the NHIS sample over a lengthy period of time implies that a large part of the sample will have moved, creating high costs of tracking and probably lowering the response rate.

4.2 Costs

The number of women to be interviewed is only one factor that influences the cost. Other factors are the number of households that have to be screened to locate the sample of women, the number that have to be tracked and visited at another address, and administrative costs related to the number of interviewers, such as travel and training.

Using a sample of persons involves no screening, except to ascertain whether the woman still lives at the address reported in the NHIS. Using the NHIS to provide an address sample involves screening the housing units selected, both the stratum of addresses which contained an eligible woman, and the subsample of other addresses. The Cycle III design also required an extensive amount of screening.

The total workload, both of interviewing and screening, is shown in Table 11.

An examination of Table 11 indicates the workloads for the Cycle III design -- both interviewing and screening -- are so much greater than for any of the competitors, that it is very unlikely that the Cycle III design will be an efficient way to proceed. A further cost analysis is necessary to distinguish between the other competitors because the sample of housing units, although requiring more interviewing and screening, does not need any tracking. The difference in cost between the 100- and 200-PSU design for the person sample also needs further examination. A much smaller interviewer force would be needed with the 100-PSU design leading to lower training and administrative costs. However, the number of interviews is slightly higher with the 100-PSU sample and more importantly, the percent of the sample to be tracked would be over twice as great.

Estimates of the direct data collection cost for each alternative are shown in Table 12, which summarizes the principal features of the alternatives considered. Only the parts of the costs that will be affected by the sample and survey design have been estimated. It is better to think of them in a relative rather than an absolute sense, since they do not include overhead, any allowance for inflation, provisions for pretests, possible changes in questionnaire content, etc. However, they do provide a basis for comparing the efficiency of the various alternatives.

On the basis of cost, the designs fall into three groups. The lowest costs are incurred with a sample of persons and one-time interviewing, with a 200-PSU design running about 10 percent higher than 100 PSU's. The next level is the housing unit sample. The costs of the three variations are almost identical. The highest costs come with the use of continuous interviewing and the Cycle III design. The Cycle III design is expensive because of the extensive screening required. Continuous interviewing costs arise because the long length of the interview period creates very small workloads on a week-to-week basis which makes it impractical to use a large staff of interviewers. The small staff, in turn, results in the use of nonresident interviewers in many of the PSU's, requiring high travel costs.

4.3 Response Rate

Table 12 also shows the expected response rates for the different plans. These estimates should also probably be viewed as a guide to the differences among procedures, rather than as estimates of the level of response rates. A somewhat lower level of response is likely for procedures that require tracking; however, the differences do not appear to be dramatic.

The information on expected costs, response rates, and interview periods shown in Table 12 provides a basis for choosing among the various alternatives.

BIBLIOGRAPHY

1. National Survey of Family Growth, Cycle III, Final Report, February 1984, Report on Methodology by Westat, Inc. for NCHS

2. Integration of Sample Design for National Survey of Family Growth Cycle IV with National Health Interview Survey, Final Report prepared by Westat, Inc. for NCHS

3. Survey Sampling 1965, Kish L., John Wiley and Sons, Inc., N.Y., p. 430.

Table 1. Required sample sizes with Cycle III design and associated sampling rates

Race and Marital Status	Sample size	Population size	Implied sampling rate 1 in __	Oversampling rate relative to white ever-married
Total	14,000	47,053,000	xx	xx
Black	6,200	7,405,000	xx	xx
Ever-married	3,600	4,157,000	1155	4.55
Never-married	2,600	3,248,000	1249	4.21
White and other races	7,800	39,648,000	xx	xx
Ever-married	5,400	28,409,000	5261	1.00
Never-married	2,400	11,239,000	4683	1.12

Table 2. Alternative survey designs

Design number	Description of design
1	Sample of person, one-time interview, 100-PSU's
2	Sample of persons, one-time interview, 200-PSU's
3	Sample of persons, continuous interview, 100-PSU's
4	Sample of persons, continuous interview, 200-PSU's
5	Sample of housing units, units without eligibles sampled at 1/2 rate
6	Sample of housing units, units without eligibles sampled at 1/3 rate
7	Sample of housing units, units without eligibles sampled at 1/4 rate
8	Cycle III model

Table 3. Sampling rates and population distribution for sampling strata used in Cycles II and III

Race and Marital Status	Sampling rate		Percent of women in strata
	Cycle II	Cycle IV	
White and Other Races			
Predominantly Black stratum	3.64r	2r	8
Other stratum	r	r	92
Black			
Predominantly Black stratum			
Ever-married	5.60r	4r	72
Never-married		5r	
Other stratum	r	r	28

Table 4. Design effects from sampling strata used for Cycles II and III

Race and Marital Status	Design effect*	
	Cycle II	Cycle IV
White and Other Races		
Ever-married	1.14	1.04
Never-married	1.14	1.04
Black		
Ever-married	1.76	1.44
Never-married	1.76	1.65

*Calculated by use of $(\Sigma p_i k_i)(\Sigma p_i/k_i)$

Table 5. Distribution of eligible women by number of eligible women in household

Number of eligible persons in households	Black		White and other	
	Ever married	Never married	Ever married	Never married
One person	75.1%	47.5%	82.6%	39.9%
Two persons				
2 ever married	1.9	-	1.3	
1 ever and 1 never married	15.6	16.6	12.3	24.1
2 never married	-	17.3	-	16.9
Three persons				
2 ever and 1 never married	1.0	0.5	0.3	0.2
1 ever and 2 never married	5.1	10.9	2.9	11.4
3 never married	-	2.5	-	3.5
Four persons*				
2 ever and 2 never married	-	-	0.1	0.2
1 ever and 3 never married	1.3	4.0	0.5	2.7
4 never married	-	0.7	-	1.1

*Five or more eligible person households have been included with four person groups.

Table 6. Mobility rates for selected population groups, 1975 to 1980.

Population Group	1 year	2 years	3 years	4 years	5 years
Total population					
Total population					
Total movers	17.7%	27.4%	35.5%	41.7%	47.0%
Local[1]	11.6	17.5	22.1	25.3	27.9
Long distance[1]	6.1	9.9	13.4	16.4	19.1
White					
Total	17.4	27.2	35.1	41.3	46.6
Local	11.1	17.0	21.4	24.3	27.1
Long distance	6.3	10.2	13.7	17.0	19.5
Black					
Total	18.3	27.3	36.2	41.9	47.2
Local	14.8	21.4	27.6	31.4	34.9
Long distance	3.5	5.9	8.6	10.5	12.3
Persons 14-44 years[2]					
Total movers	23.7	36.0	45.4	52.6	59.2
Local	15.5	20.2	28.7	32.4	35.9
Long distance	8.2	15.8	16.7	20.2	23.3
Below Poverty Level[2]					
Total	26.7	37.1	45.2	50.0	55.0
Local	18.0	24.5	28.8	31.9	34.7
Long distance	8.7	12.6	16.4	18.1	20.3
Below poverty level age 14-44[2]					
Total	35.6	48.0	56.3	61.3	68.2
Local	23.2	30.4	34.8	37.8	41.5
Long distance	12.4	17.6	21.5	23.5	26.7
Women, 16-44 years					
All women[2]					
Total	24.6	37.2	46.8	54.4	59.9
Local	16.3	24.1	29.9	34.1	36.9
Long distance	8.3	13.1	16.9	20.3	23.0
Never married[2]					
Total	20.9	27.5	37.6	43.9	47.7
Local	13.6	19.9	24.4	27.7	29.0
Long distance	7.3	7.6	13.2	16.2	18.7
Ever married[2]					
Total movers	26.2	41.3	51.1	59.8	65.7
Local	17.5	25.9	32.5	37.1	40.6
Long distance	8.7	15.4	18.6	22.7	25.1
Women 15-19 years					
Total	19.1	24.8	31.5	36.0	41.0
Local	12.5	16.6	20.5	22.6	26.1
Long distance	6.6	8.2	11.0	13.4	14.9
Total households					
Total	18.3	29.6	36.9	42.8	47.9

Note: Denominators of percentages are population totals at period after the move, and thus may differ a little from rates that would be calculated using denominators at beginning of period.

[1]Local movers are persons moving with the same SMSA, or within the same county, if outside an SMSA.

[2]Distribution between local and long distance partially estimated. Census reports do not show data on proportion of outside SMSA movers who remain in same county for these subgroups. We have estimated this by assuming the percent of local movers for total population applies to each subgroup.

Source: Bureau of the Census, Series P-20, Nos. 305, 320, 331, 353, 368.

Table 7. Design effects for Cycle IV, with Cycle III sample design

Item	Black		White and other	
	Ever married	Never married	Ever married	Never married
Diff with 14,000 sample				
P	.61	.61	.75	.75
$\bar{n}$	41	29	75	33
$\bar{\bar{n}}$	4.3	3.1	5.2	2.3
ρ_1	.0007	.0007	.0066	.0066
ρ_2	.0422	.0422	.0458	.0458
$P\bar{n}\rho_1$	.02	.01	.37	.16
$(\bar{\bar{n}}-1)\rho_2$	.14	.09	.19	.06
V_n^2	.05	.05	.05	.05
a^2	.36	.18	.25	.17
b^2	.44	.65	.04	.04
Design effect	2.01	1.98	1.90	1.48

NOTE: Design effects estimated by use of

$$\sigma^2_{x'} = \sigma^2_{SRS}[1 + a^2 + b^2 + \bar{n}\rho_1 + (\bar{\bar{n}}-1)\rho_2 + V_n^2]$$

where a^2 and b^2 are variance components arising from variability in sampling rates, ρ_1 the intraclass correlations between PSU's, ρ_2 the intraclass correlation between segments, and V_n^2 the relvariance in segment size.

Table 8. Total design effects with alternative sample designs

Sample design	Black		White and other	
	Ever married	Never married	Ever married	Never married
Current design	2.01	1.98	1.90	1.48
Alternative design				
Sample of persons with tracking				
200 PSU's				
1 year NHIS	1.44	1.24	1.44	1.27
2 year NHIS	1.50	1.28	1.52	1.30
3 year NHIS	1.56	1.33	1.60	1.33
100 PSU's				
1 year NHIS	1.44	1.24	1.54	1.31
2 year NHIS	1.50	1.28	1.62	1.34
3 year NHIS	1.56	1.33	1.70	1.37
Sample of housing units, with subsampling				
200 PSU's				
r=1/2				
1 year	1.46	1.26	1.47	1.29
2 year NHIS	1.54	1.31	1.57	1.33
3 year NHIS	1.60	1.36	1.65	1.37
r=1/3				
1 year NHIS	1.50	1.25	1.53	1.33
2 year NHIS	1.60	1.36	1.65	1.39
3 year NHIS	1.69	1.29	1.74	1.44
r=1/4				
1 year NHIS	1.55	1.33	1.58	1.38
2 year NHIS	1.67	1.41	1.74	1.45
3 year NHIS	1.74	1.48	1.84	1.51
100 PSU's				
r=1/2				
1 year NHIS	1.46	1.26	1.57	1.33
2 year NHIS	1.54	1.31	1.67	1.37
3 year NHIS	1.60	1.36	1.75	1.41
r=1/3				
1 year NHIS	1.50	1.29	1.63	1.37
2 year NHIS	1.60	1.36	1.75	1.43
3 year NHIS	1.69	1.42	1.84	1.48
r=1/4				
1 year NHIS	1.55	1.33	1.68	1.42
2 year NHIS	1.67	1.41	1.84	1.49
3 year NHIS	1.74	1.48	1.94	1.55

Table 9. Sample sizes required for equal reliability with alternative sample designs.

Sample design	Black		White and other		Total
	Ever married	Never married	Ever married	Never married	
Current design	3,600	2,600	5,400	2,400	14,000
Alternative design					
Sample of persons with tracking					
200 PSU's					
1 year NHIS	2,579	1,628	4,093	2,059	10,359
2 year NHIS	2,687	1,681	4,320	2,108	10,796
3 year NHIS	2,794	1,746	4,547	2,157	11,244
100 PSU's					
1 year NHIS	2,579	1,628	4,377	2,124	10,708
2 year NHIS	2,687	1,681	4,604	2,173	11,145
3 year NHIS	2,794	1,746	4,832	2,222	11,594
Sample of housing units with subsampling					
200 PSU's					
r=1/2					
1 year NHIS	2,615	1,655	4,178	2,092	10,540
2 year NHIS	2,758	1,720	4,462	2,157	11,097
3 year NHIS	2,866	1,786	4,689	2,222	11,563
r=1/3					
1 year NHIS	2,687	1,694	4,348	2,157	10,886
2 year NHIS	2,866	1,786	4,689	2,254	11,595
3 year NHIS	3,027	1,865	4,945	2,335	12,172
r=1/4					
1 year NHIS	2,776	1,746	4,491	2,238	11,251
2 year NHIS	2,991	1,852	4,945	2,351	12,139
3 year NHIS	3,116	1,943	5,229	2,449	12,737
100 PSU's					
r=1/2					
1 year NHIS	2,615	1,655	4,462	2,157	10,889
2 year NHIS	2,758	1,720	4,746	2,222	11,446
3 year NHIS	2,866	1,786	4,974	2,286	11,912
r=1/3					
1 year NHIS	2,687	1,694	4,633	2,222	11,236
2 year NHIS	2,866	1,786	4,974	2,319	11,945
3 year NHIS	3,027	1,865	5,229	2,400	12,521
r=1/4					
1 year NHIS	2,776	1,746	4,775	2,303	11,600
2 year NHIS	2,991	1,852	5,229	2,416	12,488
3 year NHIS	3,116	1,943	5,514	2,514	13,087

Table 10. Minimum years NHIS necessary to achieve required sample sizes and resulting sample sizes

Item	Black		White and other		Total
	Ever married	Never married	Ever married	Never married	
Minimum years NHIS					
Sample persons with tracking					
200 PSU's	1.29	1.16	.34	.54	xx
100 PSU's	2.73	2.43	.73	1.12	xx
Sample of hu's with subsampling					
200 PSU's					
r=1/2	1.46	1.27	.38	.59	xx
r=1/3	1.59	1.34	.41	.62	xx
r=1/4	1.71	1.41	.43	.65	xx
100 PSU's					
r=1/2	3+	2.75	.82	1.23	xx
r=1/3	3+	3+	.88	1.31	xx
r=1/4	3+	3+	.92	1.39	xx
Number of interviewed women					
Current design	3,600	2,600	5,400	2,400	14,000
Sample of persons with tracking					
200 PSU's	2,610	1,636	4,093	2,059	10,398
100 PSU's	2,765	1,709	4,377	2,130	10,981
Sample of hu's with subsampling					
200 PSU's*					
r=1/2	2,681	1,672	4,178	2,092	10,623
r=1/3	2,793	1,725	4,348	2,157	11,023
r=1/4	2,929	1,790	4,491	2,238	11,448

*Since a 100 PSU requires over three years accumulation of NHIS for the Black part of the sample, it is not acceptable for Cycle IV.

Table 11. Total field workload for the alternative sample design

Sample design	No. of years NHIS	Black		White and other		Total	
		Interviewed women	Screened households	Interviewed women	Screened households	Interviewed women	Screened households
Cycle III Sample Design[1]	xx	6,200	16,000	7,800	38,000	14,000	54,000
200 PSU's							
Sample of persons							
No. of years[2]	xx	1.16-1.29	xx	.34-.54	xx	xx	xx
Workload	xx	4,246	4,246	6,152	6,152	10,398	10,398
Sample of h.u.'s							
r=1/2	1.46	4,353	7,516	6,270	15,597	10,623	23,113
r=1/3	1.59	4,518	7,153	6,505	13,864	11,023	21,017
r=1/4	1.71	4,719	7,138	6,729	13,289	11,448	20,423
100 PSU's							
Sample of persons							
No. of years[2]	xx	2.43-2.73	xx	.73-1.12	xx	xx	xx
Workload	xx	4,474	4,474	6,507	6,507	10,981	10,981

[1]The number of screened households is based on the Cycle III proposal.

[2]The two numbers shown in each column are the number of years NHIS needed for ever-married and never-married women.

Table 12. Principal features of alternate sample and survey procedures

Sample design	Sample size	Households screened	Direct cost	Approximate response rate	Approximate interview periods
Sample of persons, one-time interview, 100-PSU's	10,981	10,981	$1,700,000	81 %	Oct.'87-Feb.'88
Sample of persons, one-time interview, 200-PSU's	10,398	10,398	1,870,000	82	Apr.'86-July'86
Sample of persons, continuous interview, 100-PSU's	10,981	10,981	2,630,000	83	Apr.'85-Dec.'87
Sample of persons, continuous interview, 200-PSU's	10,398	10,398	2,610,000	83	Apr.'85-June'86
Sample of housing units, units without eligibles sampled at 1/2 rate	10,623	23,113	2,020,000	84	July'86-Nov.'86
Sample of housing units, units without eligibles sampled at 1/3 rate	11,023	21,017	2,020,000	84	Aug.'86-Dec.'86
Sample of housing units, units without eligibles sampled at 1/4 rate	11,448	20,423	2,040,000	84	Sept.'86-Jan.'87
Cycle III model	14,000	54,000	2,610,000	84	xx