AN EFFECTIVE SAMPLE DESIGN FOR REACHING MINORITY POPULATIONS

Rolf M. Wulfsberg and Daniel E. Hagan, Chilton Research Services

In the continuing debate over the relative advantages of telephone interviewing versus personal interviewing, cost is frequently one factor which is cited as favoring the former technique. In studies where respondent screening is required, the cost advantages of telephone interviewing techniques can be even more pronounced. Such requirements arise frequently in marketing studies where the sponsor may wish to speak only to purchasers of a given product or in social science investigations where large numbers of minority or rare populations may be required in order to make behavioral comparisons. Such studies, where as many as eighty percent or more of the initial contacts are discarded, make personal interviewing techniques extremely expensive unless very small housing segments are used.

One such study where large numbers of minority respondents were required was the Census Exposure Study conducted by Chilton Research Services of Radnor, Pennsylvania for the United States Bureau of the Census. This evaluation study not only required that one-third of the respondents be Black and another one-third be of Hispanic origin, but also contained a requirement that large numbers of interviews be completed with each subpopulation in a short time period -- usually at most three days. The added time constraint could not even be met by telephone interviewing methods using normal Random Digit Dialing (RDD) screening techniques. To meet the severe constraints of this study, Chilton statisticians developed an innovative design which employed disproportionate sampling in telephone exchanges with higher incidences of Black and Hispanic populations. Using this approach, the incidence rates for Blacks and Hispanics were increased to 23 percent and 16 percent, respectively.

This paper describes the approach which was used, including elaborate weighting adjustment techniques which were used to minimize bias and, insofar as possible, sampling variance. The implications of this method for studies where it is necessary to oversample rare populations are also discussed.

Background of the Study and the Sample Design

The Public Information Campaign for the 1980 Decennial Census was a special program designed to reach the total population with a series of convincing, timely, and persuasive messages aimed at motivating the population toward full cooperation with the 1980 Census. "Cooperation" in this sense meant full and accurate response to the mail questionnaire or cooperative responses to the questions of an enumerator.

One requirement of the study was that Blacks and Hispanics were to be contacted in the same proportion as the White population in order to fairly assess their exposure to the Campaign. Therefore, on each day of interviewing, an attempt was made to interview an equal number of White, Black and Hispanic respondents.

The prescribed interviewing schedule and the desired sample sizes for for each of 17 waves of interviewing are shown in Table 1. An initial baseline evaluation of Census awareness was made in early February, and frequent evaluation updates were begun in March. Beginning on March 28, the minimum sample size for each wave was increased from 150 to 550.

The combined effect of an extremely short time frame for each wave of interviewing and large numbers of required completions made normal screening procedures impossible for this particular study. Instead, a scheme was developed which actually employed three separate and independent national probability telephone samples -- one aimed primarily at Whites, one aimed primarily at Blacks, and one disproportionately geared toward the identification of Hispanic households.

Table 1: Interviewing Schedule and Desired Sample Sizes

Interview Wave	Initial Contact	First Follow-Up	Second Follow-Up	Desired Sample Size 1/
1	2/4	2/5	2/6	150
2	3/3	3/4	3/5	150
3	3/10	3/11	3/12	150
4	3/17	3/18	3/19	150
5	3/19	3/20	3/21	150
6	3/21	3/22	3/23	150
7	3/24	3/25	3/26	150
8	2/25	3/26	3/27	150
9	3/26	3/27	3/28	150
10	3/28	3/29	3/30	550
11	3/31	4/1	4/2	550
12	4/2	4/3	4/4	550
13	4/3	4/4	4/5	550
14	4/7	4/8	4/9	550
15	4/11	4/12	4/13	550
16	4/14	4/15	4/16	550
17	4/18	4/19	4/20	550

1/ Completed interviews to be comprised of equal numbers of White, Black and Hispanic respondents.

Clearly, no one sample would have been optimally efficient for the conduct of the study, in that equal sample sizes of the three subpopulations were required. By constructing three independent samples, each one designed primarily for one of the three subpopulations, and combining the results of the three samples, a considerably higher level of efficiency was achieved.

Each of the three samples was a stratified two-stage cluster sample of the United States telephone household population in those counties included in the mail-out portion of the 1980 Census. For the first sample, which was directed primarily at White households, a standard Random Digit Dialing (RDD) design was employed. First, the Chilton Research Services master file of approximately 30,000 telephone exchanges was stratified by the nine Census regions. Within each region, all exchanges were assigned to a county on the basis of the principal community served. At this point, all exchanges corresponding to counties not included in the mail-out portion of the Census were eliminated. The exchanges were then sorted at the county level along three dimensions within each Census region. These dimensions were State, metropolitan/non-metropolitan status and county median income. Using a random start and systematic selection within each region -- the sampling interval was 1 out of 26 and was constant across all regions -- a total of 1,086 exchanges was selected. This procedure provided an equal probability of selection for all exchanges in the master file and assured the selection of a representative sample with respect to those dimentions mentioned above.

The two samples directed toward ethnic households involved oversampling of exchanges in counties where higher ethnic populations existed. As mentioned above, all exchanges in the master file had been assigned to a county. Data regarding the total Black and total Hispanic households in all counties in the United States were purchased from Market Statistics, Incorporated (Sales and Marketing Magazine - '79 Survey of Buying Power). These data were used to stratify the entire master file into four "county ethnic size" categories: less than 100, 100-1,000; 1,000-10,000 and over 10,000 Black or Hispanic households (this process was conducted separately for each ethnic sample).

Sampling in the smallest size category was thought to be impractical from an efficiency standpoint. Three different first stage sampling fractions were used for the selection of exchanges in the remaining three categories providing for over-representation of the counties with larger ethnic populations. All exchanges within each size category were further stratified by the nine Census regions and a systematic sample was selected using the same sampling fraction across regions.

Selection of Households Within Exchanges

The random selection of households within exchanges was accomplished for the White sample in the traditional RDD fashion. For each exchange, 10,000 possible 4-digit suffixes are available. However, local telephone companies do not assign phone numbers completely at random across these 10,000 possibilities. Banks of 100 numbers each (identified by the first two of the full 4-digit suffix) are selected for making assignments within an exchange. As a result, an average of between 50 and 60 percent of the 10,000 possibilities are not even considered when making new assignments.

To take advantage of this, a profile of which two-digit banks were open in each exchange was identified and random numbers were generated within each exchange only in those banks known to be available for residential assignments. This resulted in the elimination of over 80 percent of the non-working numbers.

The generation of the random numbers was accomplished by computer. The quantity of numbers generated in each exchange was in proportion to the size of the exchange (i.e., the percent of the 100 banks which are open). This methodology resulted in an equal probability of selection for all telephone households within a given stratum except those having multiple telephone numbers. A weighting adjustment, which is described later in this paper, was applied to multiple-telephone households to correct for the multiple opportunity of selection of such units.

Initially, a sufficient quantity of sample was generated for completion of all waves of the study. This total sample was generated in such a way so as to preserve the original stratification scheme (i.e., the sort by region, metropolitan/non-metropolitan, etc.) at the exchange level. By taking a large number of systematic subsamples of the total sample, replicates or minisamples were formed. Each replicate retains all the characteristics of the total sample and differs only in size. By taking small collections of these replicates, a sample of the desired size for each wave of the study was achieved.

The second stage sampling fraction associated with household selection for each wave of the study, therefore, has two components. The first is the sampling rate used to generate the total sample and the second is the subsampling rate used to select the replicates to complete a particular wave.

The overall sampling fraction for all telephone households in the White sample is then simply the product of the first and second stage sampling fractions.

The determination of sampling fractions and projection weights for the ethnic samples was more involved. As mentioned earlier, the first stage selection rates for the Black and Hispanic samples varied depending on the ethnic population of the county which the exchange served. While a constant second stage sampling fraction was used for all exchanges in Wave 1, the second stage sampling fraction for subsequent waves was a function of the incidence of the target population (Blacks or Hispanics) in each exchange in the sample.

Prior to Wave 2, a screening of the telephone exchanges selected for the two ethnic samples was conducted independent of interviewing. This was done to identify those exchanges which were most efficient in yielding ethnic respondents. The results of this evaluation of the sample exchanges in the Black and Hispanic samples are presented in Table 2. The percent of all sample exchanges in each sample and county ethnic size stratum achieving varying incidences of the desired ethnic households is shown. Exchanges in the two small strata exceed 15 or 20 percent incidence only occasionally. Exchanges in those counties with large ethnic populations were more productive. Those exchanges in the largest strata achieving 20 percent incidence or better for the Black sample and 15 percent incidence or better for Hispanic sample were identified for over-representation in the second stage. There were 130 such exchanges in the Black sample averaging a 49 percent incidence and 143 in the Hispanic sample averaging a 31 percent incidence.

In order to determine the projection weights for the two ethnic samples, the overall sampling fractions needed to be determined. As with the White sample, this overall rate is computed as the product of the first stage fraction and the second stage fraction. The second stage fraction in turn is a product of two rates, the first being the fraction used to generate the total sample and the second is the subsampling rate used when selecting the number of minisamples or replicates for a particular wave. The projection weight is then simply the inverse of the overall sampling fraction.

Table 2: Evaluation of Ethnic Incidence of Exchanges, by County Ethnic Size and Sample

Black Sample				Hispanic Sample			
Incidence of Black Households	County Ethnic Size 1/ Small	Medium	Large	Incidence of Hispanic Households	County Ethnic Size 1/ Small	Medium	Large
1-20%	93.9%	82.7%	75.6%	1-15%	100.0%	91.7%	72.0%
20-40	0	15.0	10.3	15-30	0	5.8	15.7
40-60	4.7	0.8	5.5	30-50	0	0	5.5
60-80	0	0	4.6	50-70	0	1.7	4.9
80+	2.3	1.5	4.0	70+	0	0.8	2.0
Sample Exchanges	45	134	525	Sample Exchanges	45	122	511

1/ Small, medium and large refer to 100-1,000 ethnic households, 1,000-10,000 ethnic households, and over 10,000 ethnic households, respectively.

Thus, three different weights are required to reflect the different sampling rates in the first wave. For Waves 2 through 17 a fourth weight was required to reflect the differential sampling at the second stage in the subset of the large ethnic strata represented by the high yield exchanges.

Because approximately equal sample sizes were desired for the three racial groups, termination of Black households had to occur prior to achieving the desired number of Hispanic interviews. Terminating all Black households contacted on Hispanic sample would have been inefficient. For this reason, Black households were interviewed for about one-third of the Hispanic sample and were terminated for the remainder of the Hispanic sample. This required the calculation of separate sets of projection weights for Black and Hispanic respondents from the Hispanic sample.

Adjustments to the Initial Weights

After the data collection phase of the study was completed, several adjustments were made to the initial weights to correct for several sample-dependent events which could not be predicted with certainty a priori. The weighting corrections included compensations for:

- o Households with multiple telephone numbers
- o Non-response
- o Differences of projected and "known" totals
- o Combining of the ethnic data from the three independent samples.

Each of these weighting adjustments is described in detail below. In order to explain the specific methodology which was employed, the following notation will be used:

t = original wave
h = sample (i.e., White, Black, Hispanic)
i = race of the respondent
j = response category (i.e., initial call, first follow-up, second follow-up, non-response)
k = household
X_{thijk} = response from (thijk) household
P_{thijk} = probablity of selection of (thijk) telephone number
W_{thijk} = inverse of probability of selection of (thijk) household, assuming one telephone per household (i.e., $W_{thijk} = 1/P_{thijk}$)
n_{thij} = sample size in (thij) cell
X_{thij} = population total for (thij) cell
d_{thijk} = number of distinct nonbusiness telephone numbers in (thijk) household

Adjustment for Multiple Telephone Numbers in a Single Household

The sample design which was employed in this study gives a predetermined probability of selection to every household telephone number in the contiguous United States. In the case of the White sample, every such number had an equal chance of selection in a given wave. In the ethnic sample, the probability of selection was the same within each incidence stratum, but the stratum probabilities differed.

While this strategy assigns fixed probabilities to household telephone numbers, it does not necessarily assign the same probability to each household. Obviously, households with more than one telephone number have multiple chances of selection. In order to adjust for this phenomenon, each respondent's initial weight was divided by the number of distinct nonbusiness telephone numbers in that household. The adjusted weights are shown below:

Initial Weight	Adjusted Weight
W_{thijk}	$W'_{thijk} = W_{thijk}/d_{thijk}$

The number of multiple household telephone numbers was ascertained in two stages. First, the number of distinct telephone numbers in the household was determined. Then, the number of those telephone numbers which are not used exclusively for business purposes was determined. Approximately seven percent of the sampled households had multiple telephone numbers and approximately five percent had multiple telephone numbers used for other than business purposes. As a result, the sum of the household weights before and after the weight adjustment described above differed by less than three percent. This phenomenon was observed in greater numbers in those households in the high ethnic incidence clusters, however. The reductions in the sum of the weights, by race/ethnicity of the respondents, are shown below:

Respondent Race/ Ethnicity	Reduction in Total Sum of Weights
White	2.4%
Black	3.1%
Hispanic	2.1%

Adjustment for Non-response

Despite the amount of time and effort applied toward the completion of interviews with a large sample of individuals, non-response is inevitable. In situations where there might be a relationship between response and variables of interest -- such as with highly sensitive topics -- moderate to high levels of non-response could introduce significant bias into the results.

In the Census Awareness Study, particularly diligent efforts were made to maximize the response rate within the design constraints. The call rule was varied according to the sample number's status after the first dialing attempt. A call placed during the daytime interviewing hours (9 AM - 4:45 PM) resulting in a no answer or unspecified callback was rescheduled for a second attempt during the evening hours. "Busy's" were dialed twenty minutes after the original call and then again in the evening if the line was still busy. Scheduled callbacks were available for a second attempt at the time specified by the interviewer.

Despite these efforts to maximize the completion rate, non-response was inevitable. The response rates were calculated in a typical fashion independently for each sample type. Completed interviews were divided by an estimate of the number of households believed to be eligible. The denominator included completes plus a percentage of the callbacks, no answer/busy and refusals. The percentage of callback and refusals estimated to be eligible for an interview was determined by the incidence rate of eligible households among those sample pieces where a determination was made. The same eligibility rate was applied to the no answer/busys after they were reduced. Based on numerous studies conducted at Chilton Research Services, it was estimated that 40 percent of the no answer/busys represented households. The remaining 60 percent are presumed to be non-assigned or businesses.

Response rates were calculated for each of the three samples for each wave of the study. These 51 response rates were evaluated after the initial day of interviewing, and again including interviews completed on subsequent days of nonresponse follow-up. Response rates for the initial day of interviewing were generally in the sixties, averaging (across all waves) 63%, 68% and 67% for the White, Black, and Hispanic samples, respectively. After including those additional interviews achieved on subsequent days, these average response rates rose to 70%, 76% and 75%, respectively.

The levels of non-response were sufficiently substantial that it was deemed necessary to adjust for the non-response. The problem of adjusting for non-response was more difficult in this study than in most due to the fact that this study contained several closely-timed interviewing waves. Because of this fact, initial non-respondents who were subsequently interviewed on later dates (referred to hereafter as "subsequent respondents") had longer time periods in which to become aware of the Census. In some cases, this longer period overlapped with the target date of a subsequent interview wave.

To reduce the potential bias resulting from the overlap problem, the data were regrouped for analysis purposes. That is, respondents whose ultimate interview date was closer to the target date of a later wave than it was to the respondent's original wave date were included with the later wave.

To account for those individuals with whom an interview was never completed, the weights for initial non-respondents who subsequently responded were increased. While this procedure still permits a potential bias, it at least bases the imputation on data from respondents who had non-respondent characteristics after the first day of the original wave.

Using the notation described earlier, the key parameters for two successive waves are shown in Figure 1. Because the regrouping of the data is quite different for various waves -- for example, in Wave 3 there was no regrouping, Wave 5 was regrouped into two analysis waves and Wave 7 was regrouped into three analysis waves -- it is difficult to represent the reweighting for all waves with a single equation. Instead, the equation will be presented for Wave t = 5. The equation for a wave such as t = 3 is the same if we let s = t instead of s = t-1. While not presented here, the equation for a wave such as t = 7 follows the same strategy of representing the non-respondents with subsequent respondents.

Figure 1: Key Parameters of Two Consecutive Waves

Wave s = t-1	Initial Respondents	Respondents to First Follow-up	Respondents to Second Follow-up	Nonrespondents
Population total	X_{shi1}	X_{shi2}	X_{shi3}	X_{shi4}
Sample size	n_{shi1}	n_{shi2}	n_{shi3}	n_{shi4}
Initial weight	W_{shi1k}	W_{shi2k}	W_{shi3k}	W_{shi4k}

Wave t	Initial Respondents	Respondents to First Follow-up	Respondents to Second Follow-up	Nonrespondents
Population total	X_{thi1}	X_{thi2}	X_{thi3}	X_{thi4}
Sample size	n_{thi1}	n_{thi2}	n_{thi3}	n_{thi4}
Initial weight	W_{thi1k}	W_{thi2k}	W_{thi3k}	W_{thi4k}

As indicated above, let t = 5, s = t-1 = 4. We wish to estimate $(X_{thi2} + X_{thi3} + X_{thi4})$. The mean for the regrouped wave of subsequent respondents is:

$$\bar{X}^{nr}_{thi.} = \left(\sum_{k=1}^{n_{thi2}} X_{thi2k} W'_{thi2k} + \sum_{k=1}^{n_{shi3}} X_{shi3k} W'_{shi3k}\right) \Big/ \left(\sum_{k=1}^{n_{thi2}} W'_{thi2k} + \sum_{k=1}^{n_{shi3}} W'_{shi3K}\right)$$

The resulting estimate of $(X_{thi2} + X_{thi3} + X_{thi4})$ is then

$$\left(\sum_{j=2}^{4} \sum_{k=1}^{n_{thij}} W'_{thijk}\right) \bar{X}^{nr}_{thi.}$$

Thus, the weights will be adjusted as shown below:

Initial Weight	Adjusted Weight
W'_{thi1k}	$W''_{thi1k} = W'_{thi1k}$
W'_{thi2k}	$W''_{thi2k} = b_{thi} W'_{thi2k}$
W'_{shi3k}	$W''_{shi3k} = b_{thi} W'_{shi3k}$

$$\text{where } b_{thi} = \left(\sum_{j=2}^{4} \sum_{k=1}^{n_{thij}} W'_{thijk}\right) \Big/ \left(\sum_{k=1}^{n_{thi2}} W'_{thi2k} + \sum_{k=1}^{n_{shi3}} W'_{shi3k}\right).$$

The effect of this adjustment is to let the $(n_{this} + n_{shi3})$ subsequent respondents carry the weight of the initial non-respondent sample of size $(n_{thi2} + n_{thi3} + n_{thi4})$. In the event that $(n_{thi2} + n_{thi3} + n_{thi4})$ is large relative to $(n_{thi2} + n_{shi3})$, this could have a very deleterious effect on the design effect and hence the sampling error. For this reason, the value of b_{thi} was not permitted to exceed 10.0. If there remained nonresponse weight which was not accounted for with $b_{thj} = 10.0$, this residual weight was distributed over the initial respondents in the appropriate wave.

Adjustment of Weights to "Known" Totals

While the previous adjustment should have reduced the potential bias attributable to nonresponse, two key problems remain relating to that adjustment. First, since computer-generated telephone numbers were used as sample, in most cases we cannot be sure that a given number corresponds to a household. Therefore, the actual "nonresponse" weight which was redistributed in the previous procedure was actually an estimate of the nonresponse weight which was assigned to households. Second, the race and ethnicity of the households which were not reached could not be determined. This means that if there were differential non-response rate among the racial/ethnic groups, the previous adjustments would not remove the accompanying biases.

One way to minimize the effects of both problems cited above is to adjust the weights for each racial/ethnic subpopulation in each wave to known United States totals if such data are available. This would also "force" the sample to appear representative, akin to the common practice often referred to as "balancing the sample". In the present study, the most recent census data (1970) were just about as out of data as they would ever be, so estimates had to be used. However, after the early 1980 Census data became available by race/ethnicity, those early counts were substituted for subsequent analyses of the data.

To demonstrate the actual adjustments which were made, again let t=5 and s=t-1=4. Define $N_{thi.}$ by

$$N_{thi.} = \sum_{j=1}^{2}\sum_{k=1}^{n_{thij}} W''_{thijk} + \sum_{k=1}^{n_{shi3}} W''_{shi3k},$$

and let N_i be the most recent census count for the number of households of race i (i = 1,2,3). Then the weights were adjusted as shown below:

Initial Weight	Adjusted Weight
W''_{thi1k}	$W'''_{thi1k} = a_{thi} W''_{thi1k}$
W''_{thi2k}	$W'''_{thi2k} = a_{thi} W''_{thi2k}$
W''_{shi3k}	$W'''_{shi3k} = a_{thi} W''_{shi3k}$,

where $a_{thi} = N_i/N_{thi.}$

Adjustment for Combined Samples

While the above weighting adjustments should have reduced most of the potential biases which have been discussed, recall that the data still were associated with three independent national probability samples. The final adjustment was done in order to permit the combining of the data from the three samples for Blacks and Hispanics. It is well known (and easily proved) that the minimum variance estimator results when the three independent estimates are weighted inversely proportional to their variance. This is, in fact, the procedure which was employed, using recent awareness of the Census as the key variable.

Let V_{thi} represent the sampling variance for race i (i = 2,3) from sample h in analysis wave t. Further, let

$$V_{hi} = (1/17) \sum_{t=1}^{17} V_{thi}.$$

Then the weight for each independent estimate is given by

$$b_{hi} = V_{1i}V_{2i}V_{3i}/[V_{hi}(V_{1i}V_{2i} + V_{1i}V_{3i} + V_{2i}V_{3i}]$$

and the final adjusted weights for i = 2,3 are as shown below:

Initial Weight	Final Weight
W'''_{thi1k}	$V_{thi1k} = b_{hi} W'''_{thi1k}$
W'''_{thi2k}	$V_{thi2k} = b_{hi} W'''_{thi2k}$
W'''_{shi3k}	$V_{shi3k} = b_{hi} W'''_{shi3k}$.

Evaluation of the Sample Design

In order to evaluate the efficiency of the sampling plan, one must examine the design effects which were introduced by that design. While the key design effect of interest in this regard is that introduced by the oversampling of high incidence minority exchanges, each adjustment to the weights affects the overall design effect. While representing somewhat of a digression from the principal topic of this paper, it is interesting from a research standpoint to examine the effect of each weighting adjustment on a key variable of interest and the associated design effects. This is discussed in the next section.

Effects of Weighting Adjustments

As described in earlier sections, five different adjustments were made to the weights before it was felt that the estimates represented minimally-biased population projections. This means that there were six different processing points at which an item and the associated design effects could be evaluated. The alphabetic notation (A through F) which will be used to discuss these six stages are defined in Figure 2.

Figure 2: Weighting Adjustment Stages

Weighting Adjustment Stage	Adjustment For: Multiple Telephones	Revised Waves	Non-Response	Known Totals	Combining Samples [1/]
A	No	No	No	No	Equal weight
B	Yes	No	No	No	Equal weight
C	Yes	Yes	No	No	Equal weight
D	Yes	Yes	Yes	No	Equal weight
E	Yes	Yes	Yes	Yes	Equal weight
F	Yes	Yes	Yes	Yes	Minimum variance

1/ For Blacks and Hispanics.

One of the most important items of interest to the U.S. Bureau of the Census was the respondents' recent awareness of the 1980 Census effort. Therefore, this item was used as the basis for the final adjustment of the weights and to examine the effects of each adjustment stage.

Estimates of the proportions of Whites, Blacks, & Hispanics who had recent awareness of the Census were evaluated for each weighting adjustment stage and each wave of the study.

For a single wave, the various adjustments to the weights occasionally caused significant shifts in the percentages. Most noteworthy in this regard was among Hispanics in Wave 9. Adjustment C caused a shift from 58.8 percent to 71.1 percent, adjustment E led to a drop from 78.4 to 60.5 percent, and adjustment F increased the percentage back to 78.9 percent. However, an examination of the various tables reveals that most of the shifts were very modest. In fact, taking the average of the percentages over all waves within each race/ethnicity group, the ultimate shifts between Stage A and Stage F were only -0.5, 1.1 and 2.6 percentage points for Whites, Blacks and Hispanics, respectively.

While the various adjustments were intended to reduce the bias in the final estimates, one must also examine the accompanying design effects. If minor reductions in bias are obtained at the expense of relatively large increases in the design effect, the mean square error of the estimate will actually increase.

As shown in Table 3, the first four adjustments to the weights tended to increase the design effects from the initial values at Stage A. This was expected since the introduction of differential weighting with the cluster level frequently has the effect of increasing the variance (and therefore the design effect). Note that the final adjustment -- combining each of the minority samples using a minimum-variance strategy -- offset the earlier increases in the design effects, however. Therefore, very minor reductions in the overall mean square errors were probably achieved for Blacks and Hispanics, though this was not the case for Whites.

Table 3: Average Design Effects, 1/ by Race of Respondent

Race of Respondent	Weight Adjustment Stage					
	A	B	C	D	E	F
White	1.14	1.14	1.14	1.47	1.47	1.47
Black	2.93	3.09	3.02	3.26	3.40	2.99
Hispanic	3.08	3.17	3.33	3.08	4.27	2.72

1/ Average design effect was calculated by averaging the design effect for each wave after deleting the two highest and two lowest design effect values.

The primary sampling specification of the study described in this paper was to complete interviews with fixed numbers of Whites, Blacks and Hispanics in a very short time frame. With respect to that objective, the multiple sample approach was very effective. The multiple sample approach increased the incidence of Blacks from 9.6 to 22.9 percent, an increase of 139 percent. Similarly, for Hispanics the increase in incidence from 2.8 to 14.9 percent represented an increase of 432 percent.

On the other hand, these increased incidences were not achieved without a cost. If pure random digit dialing (RDD) sampling had been employed without adding additional target samples, one would expect the design effects for Blacks and Hispanics to be reasonably similar to that realized among White respondents. That is, we could have expected a design effect of about 1.47 for each group, most of which would have been caused by the weighting adjustments. By adding the target samples, overall average design effects of 2.99 and 2.72 were realized for Blacks and Hispanics respectively. Therefore, the effective sample sizes in the multiple sample design were reduced by approximately 51 percent for Blacks and 46 percent for Hispanics as compared to RDD.

For general applications, cost should be considered in comparing the two methods. While the multiple sample approach uses fewer screening calls because of its increased incidences of minorities, it also requires more completed interviews to yield the same <u>effective</u> sample size (i.e., to yield the sampling variance). Since completed interviews are more expensive than screening calls, a trade-off is involved. As the length (cost) of the complete questionnaire increases relative to the length (cost) of the screening portion of the questionnaire, the multiple sample design becomes less favorable. In fact, in the current cost environment, the multiple sample approach ceases to be advantageous for interviews which are longer than 10-15 minutes.

In summary, while the multiple sample approach was an effective method for meeting very demanding specifications in the study described in this paper, it is not a superior method for all applications. For most general applications, the increased design effects will probably balance or outweigh the gains achieved with the higher incidence rates.