# EFFECT OF SAMPLING FOR NONRESPONSE FOLLOW-UP ON ESTIMATES FROM SAMPLE DATA

Eric Schindler and Alfredo Navarro[1], Bureau of the Census, Washington DC

KEY WORDS: Jackknife, Monte-Carlo, Raking

ABSTRACT: For the 2000 Decennial Census of Population only a sample of housing units which fail to return census forms by mail will be visited by enumerators. In past censuses all such households have been enumerated. In 1990 this follow-up operation required several hundred thousand temporary workers and cost over four hundred million dollars. Multiple sample designs are being considered. The objectives are (1) to reduce the cost of the census, (2) to give each person multiple opportunities to be counted, and (3) to improve quality by incorporating corrections for undercoverage into the estimates.

In 2000, as in past censuses, a sample of housing units will be asked to provide detailed housing unit, demographic, education, labor force, and income information. This paper discusses two empirical studies which examine the increase in sampling error caused by the introduction of sampling for the nonrespondents. These studies simulate a range of possible sample designs on data sets from the 1990 Census. For one of the studies it is possible to develop estimates of the between systematic sample component of the variance. The information obtained from these studies will assist in the determination of an appropriate design for the nonresponse follow-up sample which minimizes the effect on estimates from sample data.

## I. INTRODUCTION

For the 1990 Census, 60,000 weighting areas were formed from adjacent groups of blocks. The weighting areas were about the size of a Census tract, 3,000 to 6,000 persons, but were constructed slightly differently. Raking, or iterative proportional fitting, was implemented for each weighting area to insure that weighted census estimates from the long form sample matched the 100% counts for groups formed by family type, race, Hispanic origin, sex, age, and tenure. Thus, weighting areas are the basic building blocks of all long form estimation.

In 1990 and all previous censuses, data were collected from all housing units. Since about 30% of housing units did not mail back their Census forms in 1990, the field follow-up operation was very costly. In 2000, this cost will be contained by collecting data from only a sample of the nonrespondents, but only after more extensive attempts by mail. With the introduction of sampling for nonresponse follow-up (NRFU) in Census 2000, the long form sample will be changed. There will be less sample available and some long form respondents will have much higher weights. The raking procedure will no longer be matching the long form sample to 100% counts but rather to short form estimates which will also have variance. This paper examines results from two empirical studies of the effect of sampling for NRFU on estimates made from long Census form data.

Empirical Study 1 was implemented with the long form data from 480 weighting areas from the 1990 Census. These weighting areas were selected in order to overrepresent minority areas and areas where it was difficult to obtain completed long forms. Several alternative samples of nonrespondents were selected. The 1990 long form weights were adjusted to account for NRFU sampling. Mail return long forms had no changes in their weights. The 100% counts were assumed to be fixed, but the weighted NRFU sample for a subpopulation no longer matches the 100% count. This design is equivalent to collecting short form data from all nonrespondents, implementing the raking procedure as if all data were available, and then selecting long form data from a sample of long form nonrespondents. This design was expected to underestimate the standard errors.

In large counties, there are often many demographically similar weighting areas. Response problems and estimates are likely to be similar. The housing units providing long form data from 7 or 8 similar weighting areas would, by themselves, look like a weighting area, called a "pseudo-weighting area" below. The advantage of pseudo-weighting areas is that, because long form data are available for everyone, the "true" value of long form estimates can be calculated and compared to sample estimates. For 74 of the 480 weighting areas of the first empirical study pseudo-weighting areas were formed from the weighting area and seven other weighting areas in the same county with a similar proportion of minority persons.

Empirical Study 2 was implemented in the 74 pseudo-weighting areas. 1-in-6 long form samples were designated. Several alternative NRFU samples were selected. New weights were calculated for each NRFU sample and for each of 25 Jackknife samples (formed by

[1]The authors are mathematical statisticians in the Decennial Statistical Studies Division. This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.

numbering the housing units in a weighting area consecutively from 1 to 25 and dropping out one group) by the 1990 raking procedure. This study also estimates a downward bias in the 1990 error estimates because the between 1-in-6 systematic sample component of the variance, which the 1990 design did attempt to minimize, could not be estimated, but it can be here. In a simplified version of this study weights were calculated only once, after selection of the NRFU sample. This is likely to be the process in 2000.

A Monte-Carlo simulation for one pseudo-weighting area in Wisconsin provides alternative estimates of the variance to help evaluate the Jackknife variance estimates.

Section II discusses sampling, weighting, and estimation procedures. Sections III and IV cover the Jackknife and Monte-Carlo results. Section V discusses estimates of 100% characteristics made from sample data. Section VI describes some of the limitations of the studies. Section VII provides some conclusions and Section VIII provides a few references.

Thanks are due to Henry Woltman and John Thompson, among others, for their guidance and assistance over the last several years. A more comprehensive report is given in Schindler (1996).

## II. PROCEDURES

SAMPLING: Three basic designs for selecting NRFU samples from the housing units which do not respond by mail or telephone or computer during the initial phase of collection have been considered for Census 2000. The first samples directly from the nonrespondents, about 30% of the universe. An overall 1-in-6 NRFU sampling rate would require about 6 million collections in 2000, compared to 30 million in 1990. The next basic design, " 70% truncation", does full-scale NRFU enumeration in counties with less than a 70% response rate until 70% is reached A sample is selected from the remaining nonrespondents. About 7 million collections are required. The third basic design, "90% truncation", enumerates in counties until a 90% response rate is achieved. A 1-in-10 sample is selected from the remaining nonrespondents, requiring about 21 million collections, but only one-tenth of the hardest-to-collect housing units, so the savings, although less than for the other options, are still substantial. Because 90% response rates are expected to be politically more acceptable, this is the Census Bureau's current preferred design and will dominate the discussion.

Truncation at high geographic levels, such as a county, may result in low response rates in hard-to-collect neighborhoods balanced by high response rates in easy-to-collect areas. This could lead to higher standard errors for minority and renter populations and raises questions about equality of treatment if 95% of the nonminority population but only 80% or 85% of the minority population is collected. However, it would be more expensive and time-consuming to reach a 90% response rate in all smaller areas, such as tracts, before sampling. Higher sampling rates are being suggested for smaller areas with lower response rates, but even this may not be politically acceptable.

For 90% truncation, NRFU collection up to 90% was simulated with a simple random sample of housing units, because collection dates were not available. (This could bias the results of these studies because it is very hard to collect quality data from the last housing units. We have assumed that all data are good.) 1-in-3, 1-in-6, and 1-in-10 long form sampling rates were used. For the first two rates, short form sampling rates were reduced to achieve an overall 1-in-10 rate. Collecting more long forms should improve the sample data estimates, but the reduced number of short forms increases the variances of 100% data estimates. Also, a direct sample with the same number of enumerations as 90% truncation was selected with equal long and short form sampling rates.

For direct sampling and 70% truncation, higher overall NRFU sampling rates were used in the weighting areas with lower response rates. (See Navarro, 1995). Oversamples of long forms were also selected with rates dependent on the overall NRFU sampling rate. Including 100% NRFU, thirteen sets of estimates were created.

WEIGHTING: For the first empirical study, the final weights calculated by the raking procedure in 1990 were multiplied by the inverse of the NRFU sample probability. Since short forms will also be sampled in 2000, the raking procedure will have variation in both the internal cells for long form data and in the marginal cells for 100% estimates. This component of variation is not included, so the estimates are likely to be underestimates.

For the second study, the 100% estimates for the marginals of the weighting matrix for each NRFU sampling option and all 25 Jackknife samples were calculated by taking the inverse of the probability of selection into the NRFU sample. The initial weights for the interior cells were set to the inverse of the overall probability of selection in the long form sample and, for nonrespondents, the NRFU sample. In the simplified version only the full sample is weighted and Jackknife samples are formed after weighted. In either case the final weights include the uncertainty in the 100% estimates and the between systematic sample variation.

ESTIMATION: Sixteen long form characteristics were specified by the population and housing subject matter experts at the Census Bureau. Estimates of the 100% data which served as the base for sample data estimates were also created. These estimates of 100% data, based on sample data, are similar to estimates of 100% data which appear in long form data publications. They are generally not equal to the official estimates of 100% data in the published short form data tables and in 2000 they will have higher error estimates than the official 100% estimates with sampling only for NRFU.

Estimates of levels and proportions were made for the full sample and for the 25 Jackknife samples, either weighting each Jackknife sample or only the final sample. For each NRFU sampling option, standard errors for a level E or a proportion P are given by:

$$StdErr(E) = \sqrt{(1-f) \times \frac{24}{25} \times \sum_{jk=1}^{25} (\hat{E}_{jk} - \bar{E}_{jk})^2}$$

$$\text{or } StdErr(P) = \sqrt{(1-f) \times \sum_{jk=1}^{25} (\hat{P}_{jk} - \bar{P}_{jk})^2}$$

where: f is the observed sampling fraction; so 1-f is the finite population correction factor

$\hat{E}_{jk}$ is the estimate of E without the jk-th Jackknife group

$\bar{E}_{jk}$ is the average of the Jackknife estimates

$\hat{P}_{jk}$ is the estimate of P without the jk-th Jackknife group and

$\bar{P}_{jk}$ is the average of the Jackknife estimates.

For the second study, the difference between the true value and the average of the Jackknife estimates the component of variance due to the initial 1-in-6 systematic sample. This component estimating the difference between the expected value and the truth acts much like a bias. Alternative estimates of the standard error are given by:

$$StdErr'(E) = \sqrt{[(1-f) \times \frac{24}{25} \times \sum_{jk=1}^{25} (\hat{E}_{jk} - \bar{E}_{jk})^2] + [\bar{E}_{jk} - E_{true}]^2}$$

and

$$StdErr'(P) = \sqrt{[(1-f) \times \sum_{jk=1}^{25} (\hat{P}_{jk} - \bar{P}_{jk})^2] + [\bar{P}_{jk} - P_{true}]^2}$$

These formulas are probably more appropriate estimators of the error. Unfortunately, the true values will not be available in 2000, so StdErr will have to be used. Section IV shows that this may be acceptable.

## III. BASIC RESULTS

Sampling for NRFU instead of the 100% follow-up of previous censuses decreases the size of the long form sample and creates variance in the 100% counts. Both effects increase the standard errors (SE) of long form estimates. SEs and coefficients of variation (CV) depend on the characteristic and vary widely. The distribution of percentage increases in the SEs are more similar, at least for the total population (as opposed to for race/origin groups where results are often more dispersed) and will be used in the analysis across characteristics.

Table 1 on the next page shows the median values of the percent increases in the standard error between the 90% truncation NRFU sampling alternatives and the 100% NRFU alternative for all sixteen sample data characteristics in the 480 weighting areas or 74 pseudo-weighting areas. Results are shown for seven weighting options. The options which weight each Jackknife sample after NRFU sampling (cols 4-7) more closely reflect the conditions which will occur in 2000 (col 3) than those which do not (cols 1-2).

1. The results for 480 weighting areas from the first empirical study with raking before NRFU sampling

2. The results from the first study for the 74 weighting areas which were the starting blocks of the 74 pseudo-weighting areas

3. The results from the simplified version of the second study weighting only the final sample. This is probably what will be possible in 2000.

4/5. The StdErr results from two similar runs for the 74 pseudo-weighting areas without the component of variance from the systematic 1-in-6 sample.

6/7. The StdErr' results from two similar runs for the 74 pseudo-weighting areas with the component of variance due to the systematic 1-in-6 sample. Even though it may be preferable, it will not be possible to calculate these variance estimates in 2000.

There are about 7000 non-zero estimates of sample data characteristics in the 480 weighting areas and slightly more than 1000 in the 74 weighting areas and the 74 pseudo-weighting areas.

## RESULTS FOR MEDIANS:

- The median increases in the standard errors (SE) caused by sampling for NRFU are smaller for ratio adjusted 1990 weights than those for the 2000 option weighting only once which are slightly smaller than those for which each Jackknife sample is weighted.
- The results for the 74 weighting areas which were the starting points for the pseudo-weighting areas are consistent with the results for all 480 weighting areas.
- For the other weighting options, the increases in SE can be reduced from about 12-17% to about 8-14% (Columns 3, 4, and 5, rows 1,2, 3) by using higher NRFU sample rates for long forms. It is possible to put "too much" of the NRFU sample into long forms. The reduced short form NRFU sample increases the variability of the 100% data, the weighting matrix marginals, and the final weights more than the additional long form sample reduces them, thereby increasing the final SEs.
- The increases in SE caused by NRFU sampling are slightly larger for the estimator including the component of the variance due to the systematic sample.
- The last two columns of Table 1 show that the median value of the ratio of the SE with or without the

TABLE 1: **Median and 90-th Percentile Increases in Standard Error for Sample Data Estimates**

| | Long Form Sample Rate | No Reweighting (1) | (2) | 2000 Option (3) | w/o Component from Sys Sam (4) | (5) | with Component from Sys Sam (6) | (7) | $\frac{SE'_{with\ Sys\ Sam}}{SE_{w/o\ Sys\ Sam}}$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 480 WA | 74 WA | 74 PWA | 1: 74 PWA | 2: 74 PWA | 1: 74 PWA | 2: 74 PWA | 1: 74 PWA | 2: 74 PWA |
| Median Increases | High 1/3 | 5% | 6% | 8% | 12% | 14% | 13% | 13% | 1.172 | 1.178 |
| | Low 1/6 | 8% | 10% | 13% | 12% | 13% | 15% | 14% | 1.171 | 1.176 |
| | Uniform 1/10 | 9% | 10% | 12% | 15% | 15% | 17% | 16% | 1.170 | 1.181 |
| | Direct | 7% | 7% | 3% | 3% | 3% | 3% | 2% | 1.171 | 1.155 |
| 90-th Percentile Increases | High 1/3 | 36% | 40% | 35% | 56% | 55% | 61% | 62% | 1.960 | 1.957 |
| | Low 1/6 | 69% | 74% | 54% | 54% | 52% | 69% | 60% | 2.060 | 2.029 |
| | Uniform 1/10 | 103% | 107% | 61% | 64% | 61% | 70% | 72% | 1.928 | 1.974 |
| | Direct | 28% | 29% | 15% | 20% | 19% | 22% | 21% | 1.896 | 1.913 |

component of variance due to the systematic sample is between 1.15 and 1.20. The results for the 2000 option are underestimated a little more than this 15% or 20%.

- Direct sampling with the same resources decreases the median increase in the SEs from about 15% to about 3%, but requires the collection of most of the hardest-to-enumerate housing units instead of only one-tenth of them.
- The direct sampling and 70% truncation NRFU alternatives with about 7 million NRFU collections increase the SEs by 30% to 40%. 90% truncation requires three times as many collections as the other alternatives, so the increases in the SEs are smaller, about 12% to 15%. However, the final SEs are only about 20% smaller than for the smaller NRFU samples.

RESULTS FOR 90TH PERCENTILES:

It is important to examine not only the median increases in the SEs caused by NRFU sampling, but also what happens in the worst cases. Perhaps it is acceptable, given the potential savings of sampling for NRFU, to have median increases in SE of 15% but not acceptable to have some SEs doubled.

- Unless oversampling is used for long forms, the 90-th percentile of the increases in the SE error for the first empirical study are larger than the increases of the second empirical study. Weighting after NRFU sampling is more effective at controlling the high weights of outliers and the resulting variance.
- With 90% truncation, 10% of the estimates have increases of 50% or more from their 100% NRFU SEs.
- For the worst 10% of the observations, the SE with the component due to the 1-in-6 systematic sample is twice as large or larger than the SE without this component; i.e. the SE estimate in 2000 will be a 50% or more underestimate because the component due to systematic sampling is missing.
- With direct sampling or 70% truncation, 10% of the estimates have increases of 80% or more from their 100% NRFU SEs.

OTHER RESULTS:

- The proportion of persons age 5 or over who moved between 1985 and 1990 is stable, between 30% and 70%, and should have a low CV. The median CV is 6-7% for 100% NRFU and 8-9% with 90% truncation.
- The proportion of persons age 16 to 19 not in school and without high school diplomas has a small base population. The proportions are between 3% and 25%, usually under 10%. Considering the small base as well as the low proportion, the CVs should be high. The median CV, 40-50% for 100% NRFU, increases to 60-90% with 90% truncation.
- Per capita income ranges from $5,000 to $30,000. The median CV is 5-6% for 100% NRFU, but 7-8% with 90% truncation. The 90th percentile of the CV increases from 9-14% to 15-18% with NRFU sampling.
- The proportion of persons in poverty ranges from 3% to 40%. The CV is large for small values and small for large values. The median CV, 13-20% for 100% NRFU, is 20- 25% for 90% truncation.
- Thirty-four of the 74 pseudo-weighting areas were 10% or more Black. Data were extracted for these weighting areas from the second run of the second study. The increases in SE caused by sampling for NRFU (Median: 11-17%; 90th percentile: 60-80%) are only slightly larger than for the total population. The effect of including the component of the variance caused by the systematic sample is about the same as for the full population: median about 1.15, 90th percentile near 2.00.

## IV. MONTE-CARLO SIMULATIONS

One pseudo-weighting area in Wisconsin was selected for a Monte-Carlo investigation. Since the Jackknife often overestimates the standard error, it was expected that this investigation would give a more accurate estimate of the standard error and allow evaluation of the Jackknife estimates. 13.9% of the inhabitants were Black, 5.7% Asian or Pacific Islander, and 7.1% were Hispanic, allowing analysis for each group. Long form and NRFU samples were selected 129 times. SE and SE' were estimated as described above. This pseudo-weighting area has increases in the SEs due to NRFU sampling larger than the median values. Median increases above the 100% NRFU SEs of up to 20% were observed for 90% truncation. The large minority populations, coupled with a 27% poverty rate, are likely contributors to these relatively large increases.

The 129 Monte-Carlo simulations provide an alternative estimate of the SE directly from the repeated estimates:

$$SE_{M-C} = \sqrt{\frac{\sum_{i=1}^{129} (E_i - \bar{E}_{M-C})^2}{129}}$$

Using the formulation including the true value results in SEs only about 1% larger than using the average of the Monte-Carlo estimates (maximum about 12%) as opposed to the average 17% increase with the Jackknife CV estimates.

These error estimates can be compared with the Jackknife derived standard error estimates from each iteration. Averages of (Median $CV_{JK}$)/$CV_{M-C}$ and (10th%ile $CV_{JK}$)/$CV_{M-C}$ are shown in Table 2. A few of the median Jackknife CV estimates and many of the 10%-ile Jackknife CV estimates are 0 and are omitted, as is the contribution from the systematic sample.

TABLE 2: Monte-Carlo CVs over Jackknife CVs

| | Long Form Sample Rate | (Median $CV_{JK}$) / $CV_{M-C}$ Average | (10th %ile $CV_{JK}$) / $CV_{M-C}$ Average |
|---|---|---|---|
| | Overall | 1.063 | 0.686 |
| Groups | Total | 1.153 | 0.843 |
| | Black | 0.962 | 0.630 |
| | Asian | 1.046 | 0.535 |
| | Hispanic | 1.105 | 0.733 |
| | 100% NRFU | 1.097 | 0.756 |
| By Sample Design | High 1/3 | 1.050 | 0.652 |
| | Low 1/6 | 1.100 | 0.700 |
| | Unifl/10 | 1.008 | 0.698 |
| | Direct | 1.071 | 0.741 |

Overall, the Jackknife CVs without the component caused by systematic sampling are slightly larger than the Monte-Carlo CVs. The average ratio is just over 1.000. About 10% of the ratios are less than 0.700. It appears the Jackknife SE estimates even without the component from systematic sampling, and by implication for the 100% NRFU option the random group SE estimates used in 1990, are slightly conservative. The approximate 15% underestimate in the StdErr compared to StdErr' caused by omitting the component due to systematic sampling roughly balances the overestimate in StdErr' caused by using replication. It appears that StdErr would be an acceptable, perhaps even a preferable, substitute for StdErr'. The decrease of several percent in StdErr caused by weighting only the final sample and not each Jackknife sample would make the error estimates in 2000 even closer to the Monte-Carlo estimates.

On the other hand, the median of the Jackknife estimates tends to be an underestimate when estimating for a small percentage or a small population subgroup. For example, there were only an estimated 29 Black youths age 16-19, of whom an estimated six were high school dropouts. The Monte-Carlo error estimate was both larger and more reasonable than the Jackknife error estimate. For many low percentage sample estimates, the Jackknife CVs are 0 or very low, while the Monte-Carlo CVs are larger and seem more reasonable.

## V: 100% DATA CHARACTERISTICS

Published reports of sample estimates include estimates of 100% data. In order to maintain internal consistency of the sample data tables, these estimates are based on the long form sample. This insures, for example, that the published sample estimate for the number of citizens in a small area does not exceed the estimate of the total population in the same table. After 1990, a number of inquiries were received asking why a particular town had 1,234 persons in the 100% data tables but 1,243 persons in the sample data tables. These requests are not frivolous. If the estimated number of poor persons is 124, more than 10% are poor if the 1,234 from the 100% data is used, but less than 10% are with the 1,243 from the sample data. This difference could be critical in qualifying for a particular program.

With 100% NRFU the raking methodology controls the SEs of the 100% counts, producing much lower estimates of error than for otherwise comparable sample data. With NRFU sampling the raking procedure is less effective at controlling the SEs of 100% data, so much so that SEs are often doubled for large population groups.

Compared with the median 12-15% increase for sample data characteristics, the SEs with sampling for NRFU are often double the standard errors with 100% NRFU. With NRFU sampling the SEs of 100% data characteristics are comparable to the SEs for sample data characteristics instead of being much smaller. NRFU sampling eliminates much of the advantage of the raking procedure. There is no longer enough sample to force the estimates of 100% characteristics from sample data to be close to the estimates from the 100% data, which now have variance of their own.

Unlike the case for sample data characteristics, estimates for 100% data based on the sample tend to have higher CVs when differential NRFU sampling for long and short forms is used. The smaller short form samples are introducing more error into the marginals of the raking matrix. This error is carried directly into the 100% data estimates made from sample data.

As NRFU sampling increases the observed difference between the 100% data estimate and the sample data estimates, dealing with data users could become more difficult. The difference between 1,234 and 1,243 can be explained much more easily than a difference between 1,234 and 1,274.

For smaller groups defined by 100% data characteristics, such as the number of Blacks between ages 16 and 19, collapsing within the raking matrix has already eliminated the high correlation between the 100% data and the weighted sample data. The SEs based on sample data, are already about as large as those of similar sample data characteristics, and the increases induced by NRFU sampling are about the same as for long form data.

## VI: LIMITATIONS OF THE DATA

- The results of these studies were limited to areas with about 3,000 to 10,000 persons. Intermediate results from Rosenthal et al (1996) indicate that for areas with 50,000 or more persons, the median percentage increase in the (smaller) SEs is 40% to 50% compared to the approximately 15% observed at the weighting area level.
- These studies have not considered the effect that adjustment for coverage will have on the sampling errors. Coverage adjustment could further increase the level of uncertainty for sample characteristic estimates.
- In 1990, only short form data (and often of poor quality) were collected for about 10% of the designated long form housing units. This loss of sample was not modeled in the second empirical study or the Monte-Carlo study. Much of this sample loss is concentrated in the last housing units collected, so SE estimates for the NRFU options would be larger than those presented in this paper.
- The "collection" from 70% to 90% was random because the actual collection dates were not available. The data in the last 10% in this study are of higher quality that the data for the last 10% collected in 1990.
- There is a good deal of imputation in the Census, both long and short form. All data were assumed good.

## VII: SUMMARY

Based on the results of these empirical studies, it appears that for estimates from sample data:

- Sampling for NRFU increases the SE of estimates of sample data characteristics. The median value of the percent increases is 12-15% if about two-thirds of the nonrespondents are included in NRFU. The median value of the percent increases is 30-40% if approximately one-sixth of the nonrespondents are included in NRFU.
- The truncation approach, where response rates are raised to 90% before sampling, is not as efficient as collecting the same number of housing units with a direct sample. Operationally truncation requires the process of stopping collection, sampling, and restarting collection. On the other hand, truncation greatly reduces the number of hardest-to-enumerate collections.
- Selecting a larger proportion of long forms for NRFU than the overall sampling rate (and a correspondingly lower proportion of short forms) can reduce the median value of the percentage increase in the SE. It appears advisable to leave a reasonable NRFU sample allocation for the short forms.
- Analysis of the Monte-Carlo estimates for one pseudo-weighting area, indicates that, on average, the error estimates released in the past and those likely in 2000 have been or will be slightly conservative, more so for large population subgroups, less so for small subgroups.
- Final decisions concerning sampling for NRFU will be based on legal and political considerations, on cost, on the effect on short form data, and only lastly on the effect on sample data. It appears that, given the preferred design option of 90% truncation, the median effects on error estimates for sample data characteristics are reasonably small, but, for the worst 10% of the estimates, the SES are increased 50-60% by NRFU sampling.

## VIII. REFERENCES

SCHINDLER, Eric, 1993, "Files for Sample of Weighting Areas", DSSD 2000 Census Memorandum Series, O-10, Memorandum to Henry Woltman, May 3, 1993.

SCHINDLER, Eric, 1996, "Effect of Sampling for Nonresponse Follow-up on Estimates from Sample Data", internal Census Bureau draft, March 13, 1996

NAVARRO, Alfredo, "Reengineered Census: Nonresponse Follow-up (NRFU) Sampling for Tracts", DSSD 20000 Census Memorandum Series, M-5, Memorandum to Ruth Ann Killion, September 27, 1995.

ROSENTHAL, Miriam, SCHINDLER, Eric, and NAVARRO, Alfredo, 1996, "Census 2000 Sample Weighting," Proceedings of the Section for Survey Research Methods, 1996 Joint Statistical Meetings