Empirical Results of Two Sample Design Options for Sampling for the Count

by C.T. Isaki, J.H. Tsay, and Y. Thibaudeau

I. Introduction

Sampling for the count or sample census is a scenario in which a sample of units is selected from a frame and a data collection procedure is used to obtain responses from those units in the sample. For units not in sample, estimates of their characteristics are provided based on sample information. Given the time constraints imposed, two sample designs were developed for study. The two sample designs consisted of one that utilized the block as a sample unit, and another that used the mail address as a sample unit.

A. Background

The sampling for the count scenario assumes that a 1990 type list universe is to be used. Such a list provides mail information by block for Tape Address Register (TAR), Pre-list, and Pre-list pocket areas. Other areas such as Update Leave (U/L) and List Enumerate (L/E) were not "mail out" areas and were deemed out of scope for sampling purposes. The list also identifies group quarters. Since group quarters were not canvassed by mail and varied in size, they were not to be sampled for our study purposes, but were assumed to be covered 100% via other means.

The characteristics studied were those relevant to provide Voting Rights data.

For our purposes, the characteristics were race/ethnicity (Black, nonBlack Hispanic and Other) and by age (less than 18, greater than or equal to 18). The characteristics were summarized to the block, address register area (collections of blocks that approximate a tract), and the district office (DO).

B. Objectives of the Study

The objectives of the study were to develop survey designs to support the sampling for the count scenario. For each survey design we aimed to produce expected values and variances of estimators of characteristics such as total persons 18 and over (we term them 18^+ for brevity), total Black persons 18^+ etc. at the block, address register area (ARA), and district office levels. We further aimed to produce various summary statistics. For each survey design we aimed to produce statistics for several sampling rates. The resulting expected sample sizes were computed for purposes of cost estimation.

We limited sample design coverage to the "mail out" universes (TAR, Pre-list and Pre-list pocket) as defined in the 1990 Decennial Census. A separate strategy would require development to handle the L/E and U/L universes.

C. Two Survey Designs

We assumed that the DO was a reasonable basis for stratification from a data collection and processing viewpoint. Hence, both survey designs were developed assuming sampling independently between DOs. As mentioned previously, GQs were not canvassed by mail. Hence, in both survey designs, GQs were treated as being

canvassed in their entirety, i.e., there is no variability as a result of GQ estimation.

The first design developed is termed the block design. In this survey design, a stratified sample of blocks is selected and all units in the selected blocks are canvassed. The other design is termed the case or mail address unit design. In this survey design, an independent sample of cases or addresses is selected from each block in the DO. Two separate estimators of block total are considered. One estimator is the simple unbiased estimator formed by direct weighting of the sample responses. The other is a biased estimator motivated by Royall (1970). In both survey designs, initial canvassing is not necessarily done exclusively by mail. Contact is established by some other means.

The block design was selected in anticipation of a subsequent post enumeration survey (PES) similar to that in 1990 where samples of census blocks were utilized. With a block design, operation of a PES would be similar to that which was conducted in 1990. It also provides for cost savings and possibly, coverage improvement. The case design was selected because it would provide unbiased estimators of characteristics for each block. Because each block is represented by resident cases, there is a political advantage in the case design.

D. Products of the Study

The products of the study are a detailed development of two survey designs to support the sampling for the count scenario. This includes a theoretical development and properties of the survey designs proposed. Summary measures of statistics by DO were computed on a flow basis. Ideally, all 450^{+} DOs could be completed given

sufficient time. Given time limitation, 23 DOs were identified for initial processing. The 23 DOs were selected using criteria such as type of DO, racial composition, geographic location, tenure, and completion date in the 1990 census. A priority listing of the 23 DOs was followed in processing in them. We completed four DOs for this report with the understanding that if more DOs were required, we would do them at a later date. Of the four DOs processed, one had a high Black concentration, one had a high nonBlack Hispanic concentration, and one had a high nonminority concentration.

E. Discussion

The study has several limitations and does not address several issues that may be pertinent in the assessment of the sampling for the count scenario. First, the survey designs assume full response of selected sample units. Second, given that a sample is selected, the sampling for the count scenario may provide better coverage and more accurate information than the census. In our study, the sample is assumed to provide the same information. Third, no special treatment was made for non mailout list universe areas, i.e., L/E and U/L areas were assumed to be canvassed in their entirety. Finally, because of the limitation on geographic codes in the data file used, no estimates for geographic units within DOs, such as places were computed.

II. Methodology

The survey designs under study use either the block or the case (mail address) as the sampling unit. It is a condition of the sampling for the count scenario that only mail addresses as available in the 1990 census could be assumed available for use in

the construction of the survey designs that follow.

Mail addresses are available in TAR, Pre-list, and Pre-list Pocket areas. In these areas, census forms were mailed to almost all addresses in the 1990 census (when not mailed, responses were obtained via other means). In L/E and U/L areas, the census forms were not mailed out and census responses were obtained via other means. For survey design purposes, should a sampling universe included L/E and U/L areas, we assumed complete coverage. Such areas covered approximately five percent of the total population. We treated each DO as a sampling universe.

Each survey design produces block estimators of total Black 18^+, total nonBlack Hispanic 18^+, (shortened to Hispanic 18^+), total Other 18^+, and total population. Since ARAs are collections of blocks, estimators of ARA total are sums over block estimators in the ARA, thus arithmetic consistency is maintained from block to ARA to DO. Because of time considerations, we limited our characteristics to the first, second, and fourth characteristics in three DOs and added the third characteristic for the DO with high nonminority concentration.

A. Block Design

The block design consists of a stratification of blocks and a simple random sampling of blocks from within strata by DO. Certainty blocks consisting of large number of cases as well as those with ten or less cases are used. Within selected blocks, all cases are mailed census forms except for a small number of cases that are canvassed by other means. The above refers to the mailout universe. For other universes such as L/E and U/L, we assume complete canvassing by some means.

Furthermore, all persons in GQs in the DO are assumed to be canvassed completely.

One requirement for the survey design is the provision of estimates at the block level. The block design was proposed in anticipation of a PES which, in 1990, used the block as the sampling unit. A property of the block design is that it supports a 1990 type PES. Otherwise, one would consider a design that provides samples of cases in each block. This is the case design that is described in the next section.

The estimator of block total for characteristic Y under the block design is y_{ih} where

y_{ih} = Y_{ih} if the i^{th} block in stratum h is in sample

$\hat{R}_h^Y M_{ih}$ otherwise

where Y_{ih} is the total for characteristic Y in the i^{th} block in stratum h,

M_{ih} is the total number of cases in the i^{th} block in stratum h (known for all blocks) and

$\hat{R}_h^Y = \sum_i^{n_h} Y_{ih} / \sum_i^{n_h} M_{ih}$ is the ratio of the sums of n_h sample observations in stratum h.

The block estimator, y_{ih}, was motivated by the desire to use the sample observation for the block when it was selected. When it was not in sample, a ratio adjustment to its case count was used. The block estimator is biased. In fact, it can be shown that any block estimator that uses its sample observation exclusively will be biased. Apart from ratio bias, the sum of y_{ih} over blocks in the h^{th} stratum is unbiased for the total in stratum h. To reduce the effect of ratio bias, stratum sample sizes were allocated to provide a minimum of approximately ten blocks per stratum.

It can be shown that the expected value of y_{ih} is

$$E[y_{ih}] = Y_{ih}(n_h/N_h) + R_{(i)h}^Y M_{ih}(1 - n_h/N_h)$$

where $R^{y}_{(i)h} = \overline{Y}_{(i)h} / \overline{M}_{(i)h}$ and

$\overline{Y}_{(i)h}$ is the mean of N_h - 1 blocks in stratum h excluding block i. (1)

From (1) we see that the bias of y_{ih} is dictated by how close the term, $R^{y}_{(i)h} M_{ih}$, is to Y_{ih}. The variance of y_{ih}, omitting the subscript h for brevity, is

$$Var[y_{ih}] = (n/N)(1-n/N)[Y_i - R^{y}_{(i)} M_i]^2$$

$$+ n^{-1} M_i^2 \overline{M}^{-2} (1-n/(N-1))(1-n/N) \sum_{\substack{j=1 \\ \neq i}}^{N-1} (Y_j - R^{y}_{(i)} M_j)^2/(N-2) \qquad (2)$$

Under the design it is clear that the covariance between y_{ih} and $y_{kh'}$ is zero for $h \neq h'$. Otherwise an expression for the covariance between the two is exhibited in Appendix A.1. The Appendix also contains derivations of (1) and (2).

B. Case Sample Design

The case design consists of simple random without replacement sampling of cases independently from within every block except for the blocks with less than, or equal ten cases, where all cases are selected. Treatment of GQ persons, L/E and U/L areas are as handled in the block design. A constant sampling rate is applied in all other blocks. The case design allows for unbiased estimators of total characteristic for every block. Estimators of total for the ARAs and DO are formed by summing over block estimators.

Under the case design, we considered two estimators of block total. One estimator is the usual Horwitz-Thompson estimator which we call the Case 2 estimator and the other is a Royall type estimator which we term the Case 1 estimator. The

Royall type estimator, y_{ih}, for block i in stratum h is

$$y_{ih} = \sum_{j=1}^{n_{hi}} Y_{ij} + \hat{R}_h (M_{ih} - n_{hi}) \quad \text{where}$$

Y_{ij} is the Y characteristic for case j in block i

n_{hi} is the sample number of cases in block i

$$\hat{R}_h = \sum_{i}^{N_h} \sum_{j}^{n_{hi}} Y_{ij} / \sum_{i}^{N_h} n_{hi} \quad ,$$

M_{ih} is the number of cases in block i and

N_h is the number of blocks in stratum h. (3)

The subscript h denotes the h$^{\text{th}}$ stratum as defined in the block design. The blocks in the h$^{\text{th}}$ stratum are used in the second term in (3) to account for those cases not in the sample of n_{hi} cases. Under the case sample design, it can be shown that the Case 1 estimator for block i in stratum h is biased with expectation

$$E[y_{ih}] = n_{hi} \overline{Y}_{ih} + R_h (M_{ih} - n_{hi})$$

where $R_h = \sum_{i=1}^{N_h} n_{hi} \overline{Y}_{ih} / n_h \quad ,$

$$n_h = \sum_{i}^{N_h} n_{hi} \qquad (4)$$

Furthermore it can be shown that the variance of y_{ih} is

$$Var[y_{ih}] = n_{hi}^2 V\left[\sum_{j}^{n_{hi}} Y_{ijh} / n_{hi}\right] + (M_{hi} - n_{hi})^2 V(\hat{R}_h)$$

$$+ 2 n_{hi}^2 (M_{ih} - n_{hi}) n_h^{-1} V\left[\sum_{j}^{n_{hi}} Y_{ijh} / n_{hi}\right]$$

where $V(\hat{R}_h) = \sum_{i}^{N_h} (n_{hi}/n_h)^2 n_{hi}^{-1} (1 - n_{hi}/M_{ih}) S_{ih}^2,$

$n_h = \sum_{i} n_{hi}$ and

$$S_{ih}^2 = \sum_{j}^{M_{ih}} (Y_{ijh} - \overline{Y}_{ih})^2 / (M_{ih} - 1) \tag{5}$$

Similarly, the covariance between y_{ih} and y_{kh} is zero if i and k are in separate strata. When they are in the same strata, their covariance is

$$\begin{aligned} Cov(y_{ih}, y_{kh}) = {} & n_{hi} (M_{kh} - n_{hk}) \, Cov\left(\sum_{j}^{n_{hi}} Y_{ijh}/n_{hi}, \hat{R}_h\right) \\ & + (M_{ih} - n_{hi}) \, n_{hk} \, Cov\left(\sum_{j}^{n_{hk}} Y_{kjh}/n_{hk}, \hat{R}_h\right) \\ & + (M_{ih} - n_{hi}) (M_{kh} - n_{hk}) \, Var(\hat{R}_h) \end{aligned} \tag{6}$$

With the above expressions it is possible to derive the variances of block estimators, ARA estimators, and DO estimator. We note that the Case 1 estimator is unbiased for stratum totals and hence is unbiased for the DO. Because it is identically the sum of unbiased estimators for each stratum, it has the same variance for the stratum as well as the DO. For each large certainty block under the Case 1 estimation scheme, a simple unbiased estimator was used. Inclusion of the Case 1 estimator in the study was motivated by the thought that it would provide a smaller variance than the Case 2 estimator. The effect of its bias was an open question.

III. Study Description

The study was based on a 1990 Census scenario in terms of the handling of the universe of mail addresses, non mail areas, and canvassing of respondents. The study utilized 1990 Census data as well, with a modification for constructing strata and determining stratum sample sizes. The processing steps for each DO used in the study are described in Appendix A.2.

A. Data Sources - Two Files

The information available for each DO was produced by linking two source files. The first file contained a list of addresses along with relevant geographic information; the second file contained demographic information regarding the household residing at those addresses, when it applied.

B. Construction of Variables and Universe

The mail universe was defined as the set of all addresses in the file. For reasons below, we termed addresses cases. Some of these addresses may not represent housing units. Some addresses were designated as "killed," meaning the address was in the original source file, but it almost surely does not correspond to a legitimate housing unit. A weaker label was that of "delete." An address with that label was suspected not to represent a housing unit, but the evidence was not as compelling.

Independently of the killed/delete designation, each address was classified as mailable or not mailable in the 1990 Census files. Each address was subjected to a mailability test. If it passed the test, the address was deemed mailable. If it failed, an additional effort was then made to try to mend the address so as to render it mailable. The address was submitted to the test a second time. If it passed, it was accepted as mailable, otherwise it was non-mailable. Example of non-mailable addresses include those with nonexistent zip code or street-name. Non-mailable addresses were assumed to have been canvassed in some manner other than via mail.

C. Construction of Strata

Strata of blocks were constructed to support the block design and the same strata were also used in defining the Royall type estimator under the case sample design. The stratification process used in the study was based on a clustering algorithm that used four variables. The four variables used were the block's case count and the ARA's ratio of the number of Black 18^+, non Black Hispanic 18^+, and other 18^+ to the total case count. Hence, the last three variables were the same for each block in the ARA. It was felt that ARA race distributions were somewhat stable over time, and hence, the use of the 1990 distributions would provide a similar stratification result in the next census. The clustering algorithm identified certain blocks as outliers. Such blocks were treated as certainty and included in the group of blocks termed large certainty blocks.

D. Allocation of Sample to Strata / Blocks

Given the strata and a sampling rate for mail universe cases, we determined sample number of blocks per stratum and sample number of cases per block in the case design. For sample size allocation to strata of blocks, we sought to minimize the variance of the Case 2 estimator of total 18^+. The construction variable described in Appendix A.2 part III. was used. The minimization was performed subject to a constraint on total number of cases discounted for certainty blocks and limited by the sampling rate specified. No limitation was placed on GQ cases nor on L/E and U/L areas. Hence, a specified sampling rate was maintained over the mail universe. Allocation of sample sizes to stratum was forced to be not less than ten or so. This

condition was imposed to control bias and variance of the block estimator.

For the case design, we assumed a proportional allocation of sample cases to each block. Hence, for all relevant blocks, a fixed rate slightly reduced for the "≤ 10 certainty blocks" was applied in each block. This produced the obvious anomaly of some blocks being assigned smaller sample sizes than blocks of smaller size. We ignored this in the study.

E. Summary Statistics

We created block and ARA files of expected values, variances, and other statistics listed below by DO. The following summary statistics were computed.

Let

Y_i be the census total characteristic for the i^{th} area

y_i be the estimator of total characteristic for the i^{th} area

$E[y_i]$ be the expected value of y_i

$Var[y_i]$ be the variance of y_i

$Bias(y_i) = E[y_i] - [Y_i]$

$MSE(y_i) = Var[y_i] + [Bias(y_i)]^2$ be the mean square error of y_i

$Relative\ Bias(y_i) = Bias(y_i) / Y_i$

$s.e.(y_i) = Var(y_i)^{1/2}$

$RMSE(y_i) = MSE(y_i)^{1/2}$ be the root mean square error of y_i.

The summary statistics by DO and by area (block, ARA) were -

i) mean [s.e. (y_i) / E [y_i]]

ii) mean | Relative Bias (y_i) |

iii) mean [RMSE (y_i) / Y_i]

vi) Sample size count of blocks and cases for - GQ, mail out, mail universe but not mailed out, L/E and U/L

IV. Results

The sampling rates used for the first DO processed, DO 2107 - South Boston, MA, were 10, 20, 30, 40, 50, 60 and 70 percent. For the succeeding DOs, the ratios used were 10, 33 and 50 percent.

The summary statistics are organized by DO. For each DO, Tables 1 through 3 provide summary statistics for characteristics Black 18^+, Hispanic 18^+, and Total 18^+. The summary statistics in these tables are provided for sampling rates of 10, 33 and 50 percent. For each sampling rate, mean summary statistics are provided for blocks, ARAs, and the DO. The tables also provide the number of blocks and ARAs with the positive characteristic in question and summary statistics under both survey designs. Table 4 provides counts of sampling units for each sampling rate and survey design, while Table 5 provides a distribution of units for GQ, L/E and U/L universes. Tables 6 and beyond provide descriptive summaries when the true count is zero.

V. Discussion of Some Tabular Results

The following presentation discusses some censual observations in the tabular results and provides an explanation.

A. Increasing Variance of the Case 1 Estimator, as the Sampling Rate Increases.

Consider the behavior of the average relative standard error of the royall estimator, at the block level. This estimator yields an estimate of the number of individuals with a given characteristic (Black less than 18, black 18^{+}...). Typically, the relative variance increases when the sampling rate of the process increases, for small values of the sampling rate .

This behavior is attributable to the nature of the variance of the Case 1 estimator, in terms of the sampling rate. The variance is the sum of three distinct parts corresponding to the variances and the covariance between the two terms involved in the Case 1 estimator (See equation [5]). The first term is a second degree polynomial strictly increasing between 0 and .5. In the situations studied as part of this research, it was found that the first component increases for small values of the sampling rate as the sampling rate increases. As the variance of the Case 1 estimator increases with increasing sampling weight, the bias decreases. The resulting relative M.S.E. is usually decreasing, as the sampling rate increases. The table on the next page illustrates this situation with an example. For the count of Black 18^{+} in D.O. 2107, it gives the average of the variances and of their three components over all the blocks with a true count greater than zero.

Rate	1st component	2nd component	3rd component	Variance	M.S.E.
10%	5.3722	2.3105	0.4373	8.1201	1646.27
20%	9.7202	0.8695	0.4063	10.9961	588.294
30%	12.4282	0.3140	0.2578	13.0001	375.936
40%	13.7525	0.2885	0.3774	14.4184	199.916
50%	13.9300	0.1181	0.2349	14.2832	138.459
60%	12.6712	0.04638	0.1393	12.8569	78.5502
70%	10.4599	0.01968	0.0919	10.5715	42.2967

The clear pattern in this table is responsible for inducing a similar pattern for the average of the relative standard errors of the appendix.

B. Increasing Relative Bias from Block to D.O. Level (D.O. 2801)

Another interesting fact concerns the average relative bias at the ARA level for the Royall estimator compared to the average relative bias at the block level for the same estimator. For instance, the relative bias of the Case 1 estimator for Hispanics 18+ for DO 2801, is 1.1481 and at the ARA level, it is 4.3725. An ARA covers more cases than the typical block. Intuitively, the average relative bias should decrease. For the computation at the block level, the blocks with no cases of Hispanics 18^{+} are

ignored. Note however, some of these blocks without Hispanic 18^+ can have a large bias. On the other hand, For the computation at the ARA level, few ARA have no Hispanics 18^+, and all the blocks that are part of an ARA containing at least one hispanic 18^+ are included. In particular, some blocks with large bias, excluded from the block level calculation, are now included in the ARA analysis.

For example, a particular ARA (4052) in D.O. 2801, has only one Hispanic aged 18. This individual represents only one block; all the other blocks in the ARA were not part of the analysis at the block level, and their bias did not enter into the computation of the average relative bias for Hispanics 18^+, at the block level. However, some of these "zero" blocks possessed a large absolute bias. One block (120), for instance, did not have any Hispanic 18^+, but had an absolute bias of 47. In addition, several blocks with no Hispanic 18^+ also exhibited moderate-to-large absolute biases. The total absolute bias for the ARA is about 78. Since there was only one Hispanic 18^+ in this ARA, the relative bias of the ARA was 78. The relative bias of this ARA contributed heavily to the average relative bias at the ARA level.

Why block 120 in ARA (4052) possessed a large bias was found in the stratification. The rate of imputation embedded in the Case 1 estimator is a function of the stratification. At the 10% rate, the blocks were classified into four strata. Most of the blocks (77%) belong to strata 2 and 3. The associated imputation rates were .0916 and .0263. All the blocks in ARA 4052 were in stratum 3, except for block 120. Block 120 was in stratum 4 with a corresponding rate of imputation of .1349. In addition, there were 288 households in block 120. These circumstances together gave

an estimated number of hispanics 18^{+} of 47, while in fact, there were none.

C. Conflicting Trends between Absolute and Relative Measurements as the Sampling Rate Increases

The analysis of the block design also presented some unusual observations. In particular, in some cases when the sampling rate increase, the standard error decrease but the coefficient of variation increase. This phenomenon is recreated on a small scale in the two tables below, for a set of four blocks taken in D.O. 2107. These figures relate to the counts of black Americans aged 18 or more. The total for the standard error goes up when increasing the sampling rate from 50% to 60%. However, at the same time, the total for the coefficient of variation decrease. The same opposite trends can be observed for the root mean square error vs. the relative root mean square error. In the following tables, block 1 ,2, 3, 4, refer to specific blocks in D.O. 2107 (South Boston). Block 1 is in fact, block 303 in ARA 4037, block 2 is block 203 in ARA 4038, block 3 is block 305 in ARA 4038, and block 4 is block 103 in ARA 4040.

1. Sampling Rate 50%

Block #	Y	E(Y)	S.E.	C.V.	R.M.S.E.	R.M.S.E. /Y
1	27	40.8	10.357	.254	17.254	.639
2	25	27.8	5.104	.184	5.814	.233
3	3	5.0	1.384	.275	2.456	.819
4	121	131.7	9.473	.072	14.282	.118
Total			26.318	.785	39.806	1.809

2. Sampling Rate 60%

Block #	Y	E(Y)	S.E.	C.V.	R.M.S.E.	R.M.S.E./ Y
1	27	19.0	5.599	.294	9.728	.360
2	25	27.3	3.518	.129	4.223	.169
3	3	11.0	5.409	.490	9.687	3.229
4	121	131.7	8.934	.068	14.002	.116
Total			23.460	.981	37.640	3.874

VI. Acknowledgements

Several persons assisted us in our research. Sue Love (2KS) interpreted data fields and identified 1990 census data files that were used in our work. Don Dalzell and Ellen Katzoff, both of DMD, provided data files, while Maureen Lynch (SRD), converted the data files to SAS data sets. Leroy Bailey (SRD), assisted in the selection of the 23 DOs and ordered them for processing. Elizabeth Huang (SRD), constructed the noncertainty strata of blocks using a clustering algorithm.

Jim Fagan (SRD), provided optimum sample size allocations of blocks using nonlinear optimization software when the usual Neyman allocations were less than ten in a stratum.

VI. References

1. Dalzell, D. (1990), "1990 Census 21st Decennial Census 100% Detail File DCS*HEDF - < DO > draft", 13 pages.

2. Katzoff, E. and McLaughlin, G. (1992), "1990 Decennial Census - Census Data Organization Project Operation File [CDOP-OP]", 22 pages.

3. Miskura, S. (1992), Memorandum of 7/6/92 to Distribution List, Subject: 2000 Census Research and Development Alternative Designs Program with Attachments.

4. Miskura, S., Woltman, H. and Thompson, J. (1984), "Research Plan Uses of Sampling for the Census Count," Proceedings of the Social Statistics Section, pg. 458-463, Meetings of the American Statistical Association.

5. Royal, R.M. (1970), "On Finite Population Sampling Theory Under Certain Linear Regression Models", Biometrika, 57, pg. 377-387.

DO 2107

DO 2107 is South Boston, MA. It consists of 2,903 blocks of which 40 have GQ only and 2,636 have case counts that are not entirely GQ. There are 87 ARAs. The total non GQ case count is 145,393 with 9,975 not mailed but in the mail universe. The HU count from the mailed out universe of 135,418 cases was 126,127 while from the not mailed 8,159 HU's were produced. The total number of persons in the DO was 339,389. The GQ total number of persons was 14,301. There were no L/E and no U/L persons in the DO.

According to initial information DO 2107 contained 51 percent renter, 39 percent Black, 7 percent Hispanic and 63 percent of nonrespondent HU completed by 6/20/90. The percent of nonresponding HUs was 56.

DO

DO is South Boston, MA. It consists of blocks of which have case counts. There are ARAs. The total case count is with not mailed but in the mail universe. The HU count from the mailed out universe of cases was while from the not mailed HU's were produced. The total number of persons in the DO was . The GQ total number of persons was . There were L/E and U/L persons in the DO.

According to initial information DO contained percent renter, percent Block, percent Hispanic and percent of nonrespondent HU completed by 6/20/90. The percent of nonresponding HUs was .