

MEASURING THE EFFECT OF HIS PARTIALLY HIT BLOCKS ON OTHER DEMOGRAPHIC SURVEYS

Michael Ikeda, Rachel Weinstein, Todd Williams, Patrick Cantwell, Bureau of the Census*
Michael Ikeda, Statistical Research Division, Bureau of the Census, Washington, DC, 20233

Key Words: Partially hit blocks, redesign, area sampling

A. INTRODUCTION

This paper describes our research on how the National Health Interview Survey's sampling scheme affects the Census Bureau's 1990 Demographic Survey Redesign. The major surveys involved in the 1990 Redesign are the National Health Interview Survey (HIS), the Current Population Survey (CPS), the National Crime Survey (NCS), the American Housing Survey - Metropolitan Sample (AHS-MS), the Survey of Income and Program Participation (SIPP) and the Consumer Expenditure Surveys (CE).

Unlike the other surveys in the Redesign, HIS has an all-area design. Blocks partially or completely hit by HIS are placed in the area frame. Housing units in the HIS partially hit blocks have a chance of being selected by other surveys in the area frame. This results in an increase in cost for all of the non-HIS surveys in the Redesign and an increase in variances for some surveys.

Section B gives a general background on the sampling procedures of the surveys in the Redesign. Section C gives background on the problems associated with HIS sampling. Section D gives a brief outline of the methods of the simulation. The results are detailed in Section E. Section F considers briefly the implications of our results for survey costs and variances. Section G summarizes our findings. They suggest that the effect of HIS sampling on the other surveys will generally be very small. Tables and graphs following the main text illustrate many of our conclusions.

B. GENERAL BACKGROUND

Sampling for the surveys in the Redesign is done in two stages. In the first stage, primary sampling units (PSUs) are selected. PSUs are generally counties or groups of counties. PSUs in sample with certainty are called self-representing (SR) PSUs. Other PSUs are called non-self-representing (NSR) PSUs. NSR PSUs are selected with probability proportional to some measure of PSU size. In the second stage of sampling, units (housing units) are selected within each selected PSU. Our research is concerned with within-PSU sampling.

Two major restrictions affect within-PSU sampling. The first is the Title 13 confidentiality restriction. This restriction prevents the Census Bureau from releasing addresses from the decennial census address list. Second, the Office of Management and Budget (OMB) requires that no housing unit will be selected by more than one of the major surveys.

The surveys in the Redesign use two major sampling frames, which are labelled the unit frame and the area frame. The unit frame is a list of census addresses. Those blocks (blocks are roughly equivalent to ordinary city blocks) where the census addresses are easy to find are assigned to the unit frame. The remaining blocks, which tend to be in rural areas, are assigned to the area frame. The area frame is a list of blocks and their expected number of housing units.

In the unit frame, surveys sample from the census address list. Each survey, in turn, sorts the address list and then selects a systematic sample from this list. Because of OMB regulations, units selected by one survey are removed from the sample frame for the remaining surveys. The surveys may sort either units or blocks in the unit frame. CPS and NCS sort blocks. SIPP, AHS-MS and CE sort units. HIS has a unique procedure in unit frame blocks. This procedure is outlined in section C.

Each survey, in turn, sorts and selects area

frame blocks. A systematic sample is selected from the sorted block file using the expected number of housing units (adjusted for sampling by previous surveys) from the Decennial Census. When an area frame block comes into sample, an interviewer goes out into the field to list the housing units in the block. The list is shared among the surveys so that only one survey needs to list a given block even if several surveys are in the block. Specific units are assigned to sample after the listing. Because a field listing must be done, area frame sampling is more expensive than unit frame sampling.

C. PROBLEMS ASSOCIATED WITH HIS SAMPLING

The sponsor for HIS, the National Center for Health Statistics (NCHS), wishes to use the HIS sample addresses as a sampling frame for future NCHS surveys. Because of the Title 13 confidentiality restriction, this means that HIS is not allowed to sample from the census address list and therefore uses an all-area design.

OMB regulations prevent more than one survey in the Redesign from selecting the same housing unit. Because of these regulations and the Title 13 confidentiality restriction, the Census Bureau decided to place any blocks hit by HIS in the area frame.

HIS samples from the blocks in both the unit and area frames but blocks that HIS hits in the unit frame become part of the area frame. HIS sorts blocks and selects a systematic sample using the expected number of housing units from the Decennial Census.

Units in HIS partially hit blocks have a chance of being selected in the area frame by other surveys. HIS or another survey must list these partially hit blocks as they come into sample. For non-HIS surveys, this increases the listing costs for those partially hit blocks that are initially in the unit frame. Also, some surveys may realize an increase in the variance of their estimates by using a block sort (required in the area frame) rather than a unit sort. The problems of increased cost and variance would be expected to be especially acute in large cities, where many blocks are large and all or most of the PSU is initially assigned to the unit frame.

Magnifying the problems, HIS plans to "oversample" most PSUs at a rate of four times its national sampling rate. HIS wishes to increase the number of blacks and Hispanics in sample. HIS oversamples all households in the PSU, then subsamples from their selected units with differential rates of retention. Other surveys may not draw sample from any of the units selected by HIS, even the units dropped during HIS' subsampling operation. This avoids duplication while still leaving the PSU representative for surveys sampling later.

HIS assigns blocks to strata based on the percent black and percent Hispanic in the block. Units that HIS selects from strata that have a higher proportion of black or Hispanic residents have a higher probability of retention during subsampling. The stratum definitions (and the retention rate for each stratum) are the same for all PSUs. HIS-hit blocks that are initially in the unit frame are placed in the area frame even if they are dropped during HIS' subsampling operation.

D. METHODS

We used 1990 census block-level files for the New York City metropolitan area and the entire states of Illinois, Georgia, and California. With some simplifying assumptions (such as removing all blocks with no housing units) we simulated the Redesign sampling plan [1] to explore the effects of HIS sampling on the other surveys. We ran simulations for PSUs in these states and the New

York City metropolitan area where HIS and at least one other survey selected sample in the 1980 design. In all, we ran simulations on 29 PSUs.

In all simulations we oversampled HIS at its maximum rate. This is 4 times its national rate, except in Bleckley PSU in Georgia. In Bleckley PSU, the oversampling rate for HIS was reduced to 2% because of the small PSU size.

In each PSU, we included only those surveys which selected sample in that PSU in the 1980 design. To determine the survey's sampling interval (SI), we used the most current available information on 1990 sampling rates combined with the PSU's 1980 probability of selection.

We now define a key measurement that we use throughout this paper, the percent units moved (denoted by %UM). The denominator is the number of units hit by all non-HIS surveys that are initially in the unit frame. The percent units moved is the percent of these units which are in HIS partially hit blocks.

The results presented in this paper are based on the simulation run with the median value of %UM for the runs on the given PSU with the given set of surveys and sampling intervals. If there were an even number of runs, one of the two median runs was selected at random. The results are intended to be descriptive of our findings. We are not claiming statistical significance for them.

E. RESULTS AND DISCUSSION

Percent Units Moved: General.

The stem-and-leaf diagram in Figure 1 shows that %UM is less than 9% for all but two small PSUs, Bleckley in Georgia, and Madera in California. In addition to %UM, Table 3 provides blocks moved (BM). Blocks moved is the number of HIS partially hit blocks that are initially in the unit frame and are hit by other surveys.

Except for the two small PSUs mentioned, New York City had the highest value of %UM--8.54%. This is not surprising. New York City has a high proportion of large blocks. In fact, over 10% of the 28,188 (non-zero) blocks in New York City (all initially unit frame) have at least 240 housing units. In every other PSU, fewer than 5% of the blocks in the initially unit frame are this large.

The graph in Figure 2 plots %UM against the proportion of blocks with at least 240 units for the 18 PSUs that were SR in the 1980 HIS design. These values can be found in Table 4. Except for New York City, there is a strong linear relationship between these variables. See the section on "Subsegmenting Large Blocks" for the significance of 240 units.

Bleckley PSU in Georgia had the highest value of percent units moved, 28.9%, much larger than the values for other PSUs. This is a small NSR PSU with a very small HIS SI--HIS selects about 67% of the units in the PSU (NCS selects 13%).

The %UM for Madera County PSU in California was also high, 17.4%, for the same reasons as listed above: it is a small PSU with a very small HIS SI. We may have to monitor small NSR PSUs selected by both HIS and at least one other survey. However, such small PSUs usually contain only one survey. Furthermore, even if the percent units moved is high in these PSUs, only a small number of units will be moved from the unit to the area frame.

Effect of HIS Sampling Interval.

The percent units moved is strongly affected by the HIS SI. Table 1 shows five runs for New York City which differ only in the HIS SI. The %UM is roughly inversely proportional to the HIS SI.

The effect of a small HIS SI can also be seen in the high values of %UM for small PSUs such as Bleckley, Madera, Stanislaus County, and others. In Bleckley PSU, with HIS and NCS SIs as small as 18 and 54, respectively, %UM was 28.9%.

Differences Among Surveys.

We investigated whether the percent units moved for the non-HIS surveys differed systematically from each other. Tables 1, 2, and 6 show that, as expected, the individual surveys within a PSU generally have similar values of %UM. Although

the differences appear to be random, they tend to be greater in the smaller PSUs, where fewer units are moved from the unit to the area frame. Note that %UM for the individual surveys is given in the tables as %M. Units moved (UM) is also given for each survey. Units moved is the number of the given survey's initially unit-frame units which are in HIS partially hit blocks.

Variability Due to Random Start.

Generally, varying the random starting point (random start) for HIS' systematic sample has only a minor effect on percent units moved in large PSUs. In small PSUs, %UM can vary somewhat more with different random starts. This should not be a problem, however, as the actual number of units moved is small.

In small PSUs, because fewer blocks move from the unit to the area frame, different random starts can affect %UM more. In addition, small PSUs usually have fewer non-HIS surveys. This leads to less averaging of the effects of the random starts for the individual surveys and causes greater variability in %UM.

Table 5 demonstrates this result on a large and a small PSU. Varying HIS' random starts has limited effect on %UM in Los Angeles, but a larger effect in Madera County.

Blocks and Units Not Retained.

Of the HIS hits which move blocks from the unit to the area frame, some are not retained by HIS during its subsampling procedure. However, these blocks remain in the area frame.

Table 3 gives blocks not retained (BNR). Of those HIS partially hit blocks that are initially in the unit frame, blocks not retained is the number that are both hit by other surveys and not retained in the HIS subsampling process. In addition to percent units moved, we measured percent units not retained (%UNR). The denominator is the number of units hit by all non-HIS surveys that are initially in the unit frame. Percent units not retained is the percent of these units which are in HIS partially hit blocks that are not retained by HIS in its subsampling operation. By definition %UNR is between 0% and %UM. For a fixed %UM, %UNR should be higher in PSUs with mostly white residents, and lower in PSUs with a larger proportion of black or Hispanic residents. The differences are due to differential retention rates.

The value of %UNR generally does not vary much among runs within a PSU. However, within its range of 0% to %UM, %UNR is more variable in a few PSUs. These are cases with only a small number of partially hit blocks.

Table 5 gives %UNR for six different HIS random starts in Los Angeles County PSU and Madera County PSU. %UNR is more variable in Madera County, only because %UM varies more.

Non-HIS Sampling Intervals.

The SIs for surveys other than HIS do not seem to have an effect on %UM. (Of course, the number of units moved, the denominator for %UM, increases proportionally with the sampling fraction.)

Table 2 gives results for five simulations on Fresno County PSU. The five runs differ only in the SIs of the non-HIS surveys. To cover a range of probability values, we used SIs which were .5, .8, .9, 1.0 and 1.1 times the basic SIs. The basic SI, for a given survey and PSU, is the SI that would be used in a SR PSU multiplied by the probability of selecting the given PSU in the given survey. We used the 1980 design probability of selection multiplied by the national, regional, or state SI to obtain the basic SI.

The percent units moved is similar in all five runs. There does not seem to be any pattern to the minor differences that exist. This observation is supported by less extensive results from other PSUs.

Subsegmenting Large Blocks.

If the size of a block sampled in the area frame is at or above a specified cutoff, the block can be "subsegmented"--subdivided--in the field. Then only those subsegments which include units selected for sample need be listed. For large

blocks, subsegmenting can reduce some of the extra listing costs.

The subsegmenting cutoff is

- ▶ 240 units in Tape Address Register (TAR) areas,
- ▶ 160 units in Prelist areas, and
- ▶ 80 units in List/Enumerate areas. [2]

List/Enumerate areas, mostly in rural parts of the country, are automatically placed in the area frame, and so are not relevant to this research. Most large cities, our focus, are designated as TAR areas and are initially placed in the unit frame. Therefore we marked all HIS partially hit blocks which were at or above the cutoff size. The cutoff size was part of the input to the program.

As expected, large blocks are heavily overrepresented among blocks hit by HIS. For example, in New York City, HIS hit 12.99% of the 2988 blocks which have at least 240 units. On the other hand, HIS hit only 4.05% of the 25,200 (non-zero) blocks with fewer than 240 units. Subsegmenting may be a way to lessen somewhat the cost increases due to HIS sampling.

F. IMPLICATIONS FOR SURVEY COSTS AND VARIANCES

Costs.

We cannot predict with much accuracy how much HIS partially hit blocks will increase costs for other surveys. We expect that the cost per case in HIS partially hit blocks that are initially in the unit frame will be greater than that for unit frame blocks, but less than that for blocks initially in the area frame. We conclude that higher values of percent units moved should be associated with higher cost effects.

Some survey, HIS or another, will list the units moved to the area frame due to HIS sampling. However, listing a block which would have been in the unit frame--generally more urban--is expected to be cheaper than listing a block initially in the area frame. Second, urban blocks which are moved to the area frame are generally updated less frequently. This can save a large part of area frame costs.

There are other minor factors to consider. In larger cities, more large blocks are hit, allowing for more subsegmenting. But it is difficult to measure what fraction of a large block must be listed on average. Does the cost differential between the frames vary from one survey to another? Finally, in partially hit blocks which HIS does not retain during subsampling, the idea has been raised of using the Census files to do the listing. How much could we save if we did this?

The uncertainties in the available cost data prevent us from making any precise estimates of the cost increase due to HIS partially hit blocks. The following rough calculation for CPS is used to illustrate the general magnitude of cost increases we can expect. It should not be taken as a firm estimate of the actual cost increases.

In CPS, 60 households will be in sample in the 10 years of the design for any given hit string (the units in any given CPS hit). Each household will be in sample for eight interviews. Suppose that %UM for CPS is 10%. Suppose further that the entire PSU is initially assigned to the unit frame. From conversations with persons involved with the Redesign cost research a rough estimate of \$600 was obtained for the added cost resulting from shifting an entire CPS hit string to the area frame. If the added cost for an entire hit string is about \$600, then the added cost per case would be about $(.10 * \$600) / (.8 * 60)$ or about \$0.125 per case. This compares to a current total cost per case of about \$25.

It does appear that for the values of percent units moved obtained in most PSUs (including even New York City) the cost increases caused by HIS partially hit blocks should be very small.

Variance Effects.

These results do not include any quantitative estimates of the implications for the variances of key survey variables. Initially, we had planned

to investigate the effect of HIS partially hit blocks on the variances of key AHS-MS variables. AHS-MS will sort units in the unit frame, but must sort blocks in the area frame. Therefore, shifting blocks from the unit frame to the area frame reduces the efficiency of the AHS-MS sort.

Holding us back was the fact that the 1990 census files we used in this research are block files. We didn't have unit-level files to use in this research. When the 1990 unit-level census files become available, we can extend our simulation to the units within blocks, and measure variance effects.

We can say that the effect on the AHS-MS variances should generally increase with increasing percent units moved for AHS-MS. We feel that, for the values of percent units moved that were obtained for AHS-MS, the effect on variances should be very small. The effect on variances should also be very small for any other survey that wishes to sort units in the unit frame.

G. SUMMARY AND CONCLUSIONS

In the areas we sampled, the percent units moved (%UM) generally

- ▶ is small: less than 9% for all but two very small PSUs, and less than 7% for all but five of the 29 PSUs we ran,
- ▶ varies inversely with the HIS sampling interval,
- ▶ increases with the frequency of large blocks in the PSU,
- ▶ is about the same for all non-HIS surveys in a given PSU,
- ▶ does not vary with changes in the SIs of surveys other than HIS, and
- ▶ varies only slightly with a change in the random starts in large PSUs (although more so in smaller PSUs).

In each PSU we ran, at least 50% of the partially hit blocks selected by non-HIS surveys were not retained in the HIS subsampling process. This means that in more than half of the blocks where HIS sampling forces other surveys to do area listing, HIS will not interview in the block.

Large blocks (blocks above the subsegmenting cutoff) are overrepresented among blocks moved. Subsegmenting will lessen the added listing costs in these large blocks.

We did not estimate the increase in listing and updating costs due to HIS' sampling plans. It appears that the cost increases will be very small in most PSUs because of the low values we obtained for percent units moved.

We also did not try to measure the increase in variances for surveys that would prefer to sort units rather than sort blocks. Because we obtained only Census block files, only indirect methods would have worked. We do feel that any such increases in variances will be very small in most PSUs.

ACKNOWLEDGEMENTS

The authors wish to thank Lawrence Ernst, David Hubble and Dennis Schwanz for helpful comments on an earlier draft of this manuscript.

REFERENCES

- [1] Internal memorandum for P. Waite from L. Ernst, "Results of Research into the Effect of HIS Sampling in the 1990 Redesign," June 7, 1991.
- [2] Internal memorandum for WG-9 from SG-4.9, "Area Frame Subsegmenting Cutoffs," May 2, 1991.

* 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.

FIGURE 1: DISTRIBUTION OF PERCENT UNITS MOVED

Reading Figure 1

Figure 1 contains a stem-and-leaf diagram and a boxplot of Percent units moved.

Note the two unusually high values (17.4% and 28.9%) for percent units moved. These are the values for Madera County and Bleckley-Pulaski-Twiggs Counties respectively.

Stem-and-leaf diagram of Percent units moved, N=29
Leaf Unit=0.1

One observation at 28.9

```

17|4
16
15
14
13
12
11
10
9
8|5
7|68
6|2448
5|344
4|000014
3|247
2|34799
1|26
0|0
    
```

Boxplot of Percent units moved

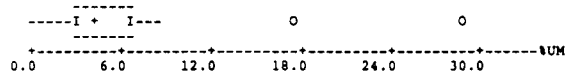


FIGURE 2: PLOT OF PERCENT UNITS MOVED VERSUS PROPORTION OF BLOCKS GREATER THAN OR EQUAL TO 240 UNITS

Reading Figure 2

Figure 2 plots Percent units moved against Proportion of Blocks greater than or equal to 240 for all HIS SR PSUs.

Proportion of blocks greater than or equal to 240 is the proportion of the non-zero initially unit frame blocks that contain at least 240 housing units.

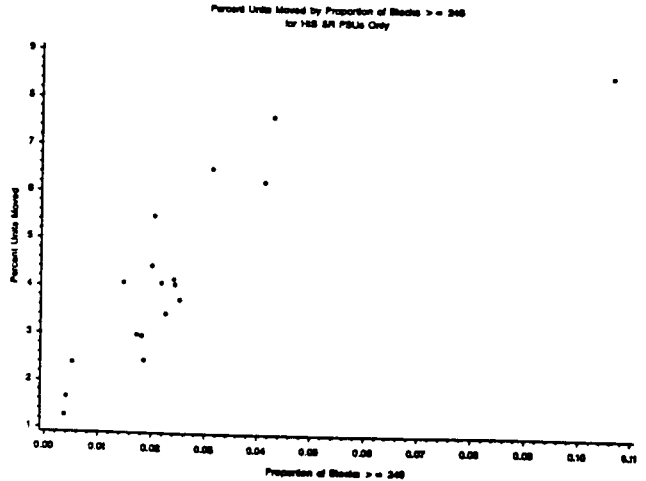


TABLE 1: DIFFERENT HIS SAMPLING INTERVALS, NEW YORK CITY

PSU							NUM	NUMR
New York City							16.69	9.89
	HIS	CPS	NCS	AHS	SIPP	CE		
SI	200	1420	1658	798	1400	2138		
UM		7392	2100	628	5696	8424		
WM		16.70	16.63	16.74	16.66	16.71		
PSU							NUM	NUMR
New York City							8.54	4.89
	HIS	CPS	NCS	AHS	SIPP	CE		
SI	400	1420	1658	798	1400	2138		
UM		3780	1064	320	2912	4320		
WM		8.54	8.44	8.53	8.52	8.57		
PSU							NUM	NUMR
New York City							6.52	3.99
	HIS	CPS	NCS	AHS	SIPP	CE		
SI	533	1420	1658	798	1400	2138		
UM		2856	812	248	2240	3312		
WM		6.46	6.44	6.61	6.55	6.57		
PSU							NUM	NUMR
New York City							4.28	2.40
	HIS	CPS	NCS	AHS	SIPP	CE		
SI	800	1420	1658	798	1400	2138		
UM		1848	532	164	1472	2196		
WM		4.18	4.22	4.37	4.31	4.36		
PSU							NUM	NUMR
New York City							2.09	1.39
	HIS	CPS	NCS	AHS	SIPP	CE		
SI	1600	1420	1658	798	1400	2138		
UM		924	252	80	704	1080		
WM		2.09	2.0	2.13	2.06	2.14		

TABLE 2: DIFFERENT NON-HIS SAMPLING INTERVALS, FRESNO

PSU						NUM	NUMR
Fresno: Fresno County						5.6	3.2
	HIS	CPS	NCS	SIPP	CE		
SI	236	1329	829	700	468		
UM		172	84	250	903		
WM		5.4	4.9	5.4	5.8		
PSU						NUM	NUMR
Fresno: Fresno County						6.1	4.1
	HIS	CPS	NCS	SIPP	CE		
SI	236	2126.4	1326.4	1120	748.8		
UM		168	56	192	541		
WM		8.3	5.3	6.6	5.6		
PSU						NUM	NUMR
Fresno: Fresno County						5.6	2.9
	HIS	CPS	NCS	SIPP	CE		
SI	236	2392.2	1492.2	1260	842.4		
UM		84	56	129	514		
WM		4.8	5.9	5.1	6.0		
PSU						NUM	NUMR
Fresno: Fresno County						5.5	3.1
	HIS	CPS	NCS	SIPP	CE		
SI	236	2658	1658	1400	936		
UM		84	83	104	412		
WM		5.3	9.3	4.6	5.4		
PSU						NUM	NUMR
Fresno: Fresno County						5.6	2.4
	HIS	CPS	NCS	SIPP	CE		
SI	236	2923.8	1823.8	1540	1029.6		
UM		84	28	96	426		
WM		5.9	3.7	4.6	6.0		

TABLE 3: PERCENT UNITS MOVED AND PERCENT UNITS NOT RETAINED

NEW YORK PSUS

PSU	HIS SI	BM	BNR	%UM	%UNR
New York City	400	318	183	8.54	4.89
PSU	HIS SI	BM	BNR	%UM	%UNR
Rockland and Westchester Co.	400	28	22	2.98	2.59
PSU	HIS SI	BM	BNR	%UM	%UNR
Orange County	400	3	3	1.27	1.27
PSU	HIS SI	BM	BNR	%UM	%UNR
Nassau and Suffolk Co.	400	38	30	1.66	1.38

ILLINOIS PSUS

PSU	HIS SI	BM	BNR	%UM	%UNR
Chicago: Cook+6 more Co.	400	214	160	4.07	2.99
PSU	HIS SI	BM	BNR	%UM	%UNR
Clinton, Jersey+3 more Co.	400	14	9	2.39	1.44
PSU	HIS SI	BM	BNR	%UM	%UNR
Peoria, Tazewell, Woodford Co.	162	9	7	6.82	5.68
PSU	HIS SI	BM	BNR	%UM	%UNR
Boone and Winnebago Co.	122	21	13	7.80	4.51
PSU	HIS SI	BM	BNR	%UM	%UNR
La Salle County	47	2	1	3.23	0.23

GEORGIA PSUS

PSU	HIS SI	BM	BNR	%UM	%UNR
Atlanta	400	119	80	6.49	3.92
PSU	HIS SI	BM	BNR	%UM	%UNR
Dade, Walker, Catoosa Co.	208	1	1	4.02	4.02
PSU	HIS SI	BM	BNR	%UM	%UNR
Columbia, McDuffie, Richmond	163	0	0	0	0
PSU	HIS SI	BM	BNR	%UM	%UNR
Blackley, Pulaski, Twiggs Co.	18	19	12	28.9	19.5

CALIFORNIA PSUS: ALL SURVEYS PRESENT

PSU	HIS SI	BM	BNR	%UM	%UNR
Los Angeles	400	201	125	3.7	2.35
PSU	HIS SI	BM	BNR	%UM	%UNR
Oakland	400	78	51	4.1	3.21
PSU	HIS SI	BM	BNR	%UM	%UNR
San Francisco	400	64	44	3.4	2.55
PSU	HIS SI	BM	BNR	%UM	%UNR
Riverside, San Bernardino Co.	400	80	47	4.4	2.21
PSU	HIS SI	BM	BNR	%UM	%UNR
San Jose: Santa Clara Co.	400	50	34	4.0	2.86
PSU	HIS SI	BM	BNR	%UM	%UNR
San Diego	400	114	80	7.6	5.38
PSU	HIS SI	BM	BNR	%UM	%UNR
Orange County	400	98	67	6.2	4.56

CALIFORNIA PSUS: NOT ALL SURVEYS PRESENT

PSU	HIS SI	BM	BNR	%UM	%UNR
Sacramento	400	35	28	5.3	4.53
PSU	HIS SI	BM	BNR	%UM	%UNR
Fresno	236	23	13	5.5	3.06
PSU	HIS SI	BM	BNR	%UM	%UNR
Bakersfield: Kern Co.	185	8	4	2.7	1.39
PSU	HIS SI	BM	BNR	%UM	%UNR
Stockton: San Joaquin Co.	237	10	8	4.1	3.53
PSU	HIS SI	BM	BNR	%UM	%UNR
Napa and Solano Co.	400	10	6	5.5	3.10
PSU	HIS SI	BM	BNR	%UM	%UNR
Ventura County	400	9	6	2.5	2.06
PSU	HIS SI	BM	BNR	%UM	%UNR
Modesto: Stanislaus Co.	141	5	3	6.5	4.46
PSU	HIS SI	BM	BNR	%UM	%UNR
Sonoma County	400	2	1	3.0	1.48
PSU	HIS SI	BM	BNR	%UM	%UNR
Madera County	42	24	16	17.4	10.2

TABLE 4: FREQUENCY OF LARGE BLOCKS AND PERCENT UNITS MOVED, HIS SR PSUS

PSU	INT UFB	#B>=240	%B>=240	%UM
New York City	28188	2988	10.60	8.54
Rockland-Westchester	11064	189	1.71	2.98
Orange County (NY)	3006	11	0.37	1.27
Nassau and Suffolk Counties	34293	137	0.40	1.66
Chicago	69442	1015	1.46	4.07
Clinton, Jersey + 3 Counties	9607	49	0.51	2.39
Atlanta	22381	696	3.11	6.49
Los Angeles	58797	1473	2.51	3.72
Oakland	17373	416	2.39	4.15
San Francisco	12651	284	2.24	3.43
Riverside-San Bernardino	19325	383	1.98	4.43
San Jose: Santa Clara Cty.	11518	278	2.41	4.04
San Diego	15408	657	4.26	7.61
Orange County (CA)	15204	624	4.10	6.23
Sacramento	11449	248	2.17	4.07
Napa and Solano Counties	3656	74	2.02	5.48
Ventura County	5093	94	1.85	2.45
Sonoma County	3218	58	1.80	2.96

INT UFB is the number of non-zero blocks that are initially unit frame.

#B>=240 is the number of blocks with at least 240 housing units that are initially unit frame.

%B>=240 is the proportion of the non-zero initially unit frame blocks with at least 240 housing units.

TABLE 5: DIFFERENT RANDOM STARTS, LOS ANGELES AND MADERA COUNTIES

PSU						
Los Angeles: Los Angeles County						
	HIS	CPS	NCS	AHS	SIPP	CE
SI	400	1544	1658	724	1400	2225
HIS Random Start						
	%UM	%UNR	%BNR			
.0211	3.8	2.5	67			
.1301	3.5	2.3	62			
.4453	3.6	2.2	60			
.6082	3.7	2.3	62			
.7734	3.9	2.2	55			
.9797	3.7	2.5	65			
PSU						
Madera: Madera County						
	HIS	NCS	SIPP			
SI	42	199	200			
HIS Random Start						
	%UM	%UNR	%BNR			
.0253	14.3	9.3	68			
.1834	14.2	8.4	56			
.3301	24.3	14.1	66			
.5840	16.5	10.1	69			
.7734	17.4	10.3	67			
.9165	20.8	14.1	76			

The HIS random start is expressed as a proportion of the HIS sampling interval.

%BNR: the denominator for %BNR is the number of initially unit frame HIS partially hit blocks that are hit by other surveys. %BNR is the percent of these blocks that are not retained during HIS subsampling.

TABLE 6: CONDENSED OUTPUT TABLES FOR HIS SR PSUS

NEW YORK PSUS							
PSU						%UM	%UNR
NY City: Bronx, Kings, NY + 2 more Co.						8.54	4.89
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	1420	1658	798	1400	2138	
UM		3780	1064	320	2912	4320	
WM		8.54	8.44	8.53	8.52	8.57	
PSU						%UM	%UNR
NY: Rockland and Westchester Co.						2.98	2.59
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	1537	1658	798	1400	2423	
UM		168	56	20	176	128	
WM		3.03	3.28	3.89	3.76	2.15	
PSU						%UM	%UNR
NY: Orange Co.						1.27	1.27
	HIS	CPS	CE				
SI	400	1537	2423				
UM		2	22				
WM		.22	2.30				
PSU						%UM	%UNR
NY: Nassau and Suffolk Counties						1.66	1.38
	HIS	CPS	NCS	SIPP	CE		
SI	400	1537	1658	1400	2423		
UM		168	56	160	252		
WM		1.42	1.53	1.61	1.95		

GEORGIA PSUS

Atlanta: DeKalb, Forsyth + 16 more Co.							
PSU						%UM	%UNR
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	3314	1658	203	1400	1796	
UM		342	282	317	659	1321	
WM		5.65	6.94	6.70	6.04	6.85	

CALIFORNIA PSUS: NOT ALL SURVEYS PRESENT

Sacramento: Placer, Sacramento, Yolo Ctys.							
PSU						%UM	%UNR
	HIS	CPS	NCS	SIPP	CE		
SI	400	2658	1658	1400	1919		
UM		168	140	256	509		
WM		4.4	6.8	4.7	5.6		
PSU						%UM	%UNR
Napa and Solano Counties						5.5	3.10
	HIS	CPS	SIPP	CE			
SI	400	2658	1025	2965			
UM		55	165	72			
WM		4.5	7.1	4.1			
PSU						%UM	%UNR
Ventura County						2.5	2.06
	HIS	CPS	NCS	SIPP	CE		
SI	400	2658	1658	1400	3572		
UM		0	28	32	108		
WM		0	3.2	1.4	5.2		
PSU						%UM	%UNR
Sonoma County						3.0	1.48
	HIS	CPS	CE				
SI	400	2658	2965				
UM		0	72				
WM		0	4.8				

ILLINOIS PSUS							
Chicago: Cook, Du Page, Kane + 4 more Co.						%UM	%UNR
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	2100	1658	466	1400	2329	
UM		1092	448	255	1262	1753	
WM		4.0	3.89	4.34	4.04	4.15	
PSU						%UM	%UNR
IL MSA: Clinton, Jersey + 3 more Co.						2.39	1.44
	HIS	CPS	NCS	SIPP	CE		
SI	400	1726	1658	1400	1598		
UM		84	28	53	64		
WM		3.57	3.45	2.44	1.51		

CALIFORNIA PSUS: ALL SURVEYS PRESENT

Los Angeles: Los Angeles County							
PSU						%UM	%UNR
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	1544	1658	724	1400	2225	
UM		1518	532	160	1328	1877	
WM		3.6	4.0	3.7	3.7	3.7	
PSU						%UM	%UNR
Oakland: Alameda and Contra Costa Counties						4.1	3.21
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	2658	1658	156	1400	2965	
UM		252	140	204	402	411	
WM		3.9	4.1	4.0	4.3	4.2	
PSU						%UM	%UNR
San Francisco: Marin, S.F., S. Mateo Ctys.						3.4	2.55
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	2658	1658	131	1400	2965	
UM		84	112	204	272	324	
WM		1.6	4.0	4.0	3.5	3.9	
PSU						%UM	%UNR
Riverside and San Bernardino Counties						4.4	2.21
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	2658	1658	99	1400	3572	
UM		336	84	349	430	262	
WM		5.6	2.7	4.6	5.0	3.5	
PSU						%UM	%UNR
San Jose: Santa Clara County						4.0	2.86
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	2658	1658	97	1400	2965	
UM		168	84	232	288	213	
WM		4.0	3.8	4.2	4.7	3.3	
PSU						%UM	%UNR
San Diego: San Diego County						7.6	5.38
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	2658	1658	184	1400	1592	
UM		672	252	364	774	1528	
WM		9.1	6.6	7.3	7.4	7.4	
PSU						%UM	%UNR
Orange County						6.2	4.56
	HIS	CPS	NCS	AHS	SIPP	CE	
SI	400	2658	1658	203	1400	3572	
UM		420	224	252	608	576	
WM		6.2	6.1	5.9	6.1	6.6	