UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

March 30,199

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R- 5

MEMORANDUM FOR Dennis Stoudt
Assistant Division Chief, Processing Systems
Decennial Systems and Contracts Management Office

From: Donna Kostanich DK
Assistant Division Chief, Sampling and Estimation
Decennial Statistical Studies Division

Prepared by: Matt Salganik MS and Randy ZuWallack RZ
Decennial Statistical Studies Division

Subject: Accuracy and Coverage Evaluation (ACE) Survey: Universe File and Block Cluster Sampling Parameter File Specification

I. Introduction

This specification describes how to create the Universe File and the Block Cluster Sampling Parameter File for the selection of the initial ACE sample clusters. The Universe File contains sampling information for all the 2000 block clusters. The initial ACE sample block clusters will be selected from the Universe File using the Block Cluster Sampling Parameter File. One Universe File will be created for each of the 50 states, the District of Columbia, and Puerto Rico. The resulting Universe and Block Cluster Sampling Parameter Files will be input to the initial block cluster sampling operation which is specified separately.

The initial sample of block clusters is allocated to states according to the previously planned Integrated Coverage Measurement (ICM) 750,000 housing unit design. This sample will be provided to the Field Division for independent listing. The results from the independent listing will be used to select a reduced sample for ACE of approximately 300,000 housing units. Plans are underway to redesign the ICM into the ACE. Requirements and details of the ACE design are not known at this time.

Before the ACE universe can be created for each state, the block clustering operation must be completed and approved. The Geography Division (GEO) is producing block clusters for each state on a flow-basis. As the Decennial Statistical Studies Division (DSSD) reviews the block clustering for each state, the DSSD will notify the GEO that a state has been approved, and the GEO will deliver the Block Cluster File for the state to the Decennial Systems and Contracts Management Office (DSCMO). This state flowing process will continue through the creation of

the ACE Universe File and Block Cluster Sampling Parameter File. After the DSSD reviews and approves the Universe File and Block Cluster Sampling Parameter File for a state, the block cluster sampling will occur for that state.

These specifications should be used to flowchart the process, to generate further discussion on requirements, to identify and finalize the record layouts of input and output files, and to write computer software to implement the methodology. During and after a testing phase, it is likely that changes to the specifications will be necessary.

The sections of this specification are ordered as follows:

- The second section states the assumptions and definitions of the file creation process.
- The third section describes the creation of the Universe File.
- The fourth section describes the creation of the Block Cluster Sampling Parameter File.
- The fifth section describes the inputs into the entire process.
- The sixth section describes the outputs of the entire process.

Any comments or questions should be directed to Randy ZuWallack (x1963), Matt Salganik (x3636), or Debbie Fenstermaker (x4195).

II. Assumptions/Definitions

A. All 50 states plus the District of Columbia will use the same set of criteria for assigning block clusters to the demographic and tenure groups.

B. Block clusters are classified into three size categories based on the number of housing units: small (0 to 2), medium (3 to 79), and large (80+).

C. A separate sample of small block clusters will be drawn from each state.

D. Large block clusters will be selected at a higher rate than mediums to complement a subsequent housing unit subsampling operation in large blocks.

E. A separate American Indian Reservation (AIR) sample will be selected.

F. Remote Alaska will not be part of the ACE universe and therefore will not be sampled.

III. Creating the Universe File

The Universe File will be a 2000 collection block level file with cluster level information for sorting and sampling. It will also be the source of information for calculating the sample parameters in the next section of this specification. A separate Universe File will be created for

each state. These Universe Files will be the input into the block cluster sampling operation specified in a different memorandum.

Building the Universe File requires two major operations. First, the data from the 1990 census is merged with the 2000 collection blocks. Then each block cluster is categorized into several groups which will facilitate assigning clusters to sampling strata and sorting before sample selection. Process each state separately as specified below.

A. Assigning 1990 Data to 2000 Collection Blocks

The purpose of this section of the specification is to explain the process for using information about the 1990 tabulation blocks to estimate the characteristics of 2000 collection blocks.

The relationships between the 1990 tabulation blocks and the 2000 collection blocks are complex. In some cases one 1990 tabulation block is associated with many 2000 collection blocks. In some cases it is the other way around -- several 1990 tabulation blocks are associated with one 2000 collection block. It is also possible for one 1990 tabulation block to be associated with one and only one 2000 collection block. The information about these associations is contained in the Block Equivalency Files that the GEO has created.

The method we will use accounts for these changes in the block definitions and proportionally assigns the 1990 population and housing unit (HU) counts to the 2000 collection blocks. It will also assign the 1990 Urban/Rural and Urban Area Size Codes to the 2000 collection blocks. This method is detailed below and is accompanied by an example.

The next steps will create the Merged Data File which is a collection of the data needed to create the Universe Files.

1. Gathering 1990 Census Data

The basis for this operation is the Block Equivalency Files (see Attachment A for a layout). The GEO created one file for each county. Each record on the Block Equivalency File represents a 1990 tabulation block/2000 collection block association. The following steps place data from the Block Cluster File and the 1990 Hundred Percent Edited Detail File (HEDF) onto the Block Equivalency File.

a. For each state, append the county Block Equivalency Files together to create a state file. From this point on, this file will be referred to as the Merged Data File.

b. Sort the Merged Data File by county and 2000 collection block number.

c. Sort the Block Cluster File (for a layout see Attachment B) by county and 2000 collection block number.

d. For each 2000 collection block in the Merged Data File, obtain the HU count and geography cluster number for that block from the Block Cluster File. There are several HU counts on the Block Cluster File. Use the count found in the 'Number of HUs in Block' field. Append this count and the cluster number to every record involving the 2000 collection block in the Merged Data File. It is important to note that each collection block may have several records because of the nature of the Merged Data File. Each appearance of a collection block should have an HU count and cluster number appended to it.

   Because Remote Alaska will not be sampled, it will not be included on the Block Cluster File. Similarly, water blocks will not be on the Block Cluster File. However, both of these block types will be included on the Merged Data File. Merging the two files will result in records containing blank values. Delete these records from the Merged Data File.

e. Sort the blocks within the Merged Data File by county, 1990 tract, and 1990 tabulation block number and suffix.

f. For each 1990 tabulation block plus suffix, use race, Hispanic origin, and tenure variables on the 1990 HEDF to tally the number of people living in regular HUs in the following 12 demographic/tenure groups. Do not include people living in Group Quarters (GQ). See Attachment C for definitions of each group. Demographic information does not exist for Puerto Rico, so set

these 12 values to zero. Append these values to every record corresponding to the 1990 tabulation block plus suffix in the Merged Data File.

Hawaiian and Pacific Islander Renter
Hawaiian and Pacific Islander Owner
American Indian and Alaska Native Renter
American Indian and Alaska Native Owner
Asian Renter
Asian Owner
Hispanic Renter
Hispanic Owner
Black Renter
Black Owner
White and Other Renter
White and Other Owner

Hispanics can be of any race, however we want these groups to be mutually exclusive. So, place all Hispanics in the Hispanic group regardless of race.[1] For example, if a person is a Hispanic Asian Renter, then she is classified as a Hispanic Renter in this process

g. Calculate the total number of people living in owned HUs and total number of people living in rented HUs in each 1990 tabulation block plus suffix. These two values can be calculated for Puerto Rico using the H4 variable on the HEDF (see Attachment C for more information). Append these values to every record corresponding to the 1990 tabulation block plus suffix in the Merged Data File.

Total Renters
Total Owners

h. Calculate the total number of regular HUs and the total number of occupied regular HUs in each 1990 tabulation block plus suffix (see Attachment C for definitions). This information may be used later by the Long Form Sample Design and Estimation Team. Do

[1]Note that Hispanic American Indians and Hispanic Alaska Natives living on American Indian reservations, trust lands, tribal jurisdiction statistical areas, tribal designated statistical areas, and Alaska native village statistical areas may be classified as American Indian and Alaska Native for estimation purposes. For sampling purposes, we will treat Hispanic American Indians and Hispanic Alaska Natives as Hispanic.

not include any people living in GQs in any of these calculations. Append these values to every record corresponding to the 1990 tabulation block plus suffix in the Merged Data File.

Total Housing Units
Occupied Housing Units

i. Calculate the total number of people living in HUs in each 1990 tabulation block plus suffix. Append this value to every record corresponding to the 1990 tabulation block plus suffix in the Merged Data File.

Total People (Non-GQ)

j. From the 1990 HEDF, get the geography codes listed below for each 1990 tabulation block plus suffix. For 1990 tabulation blocks containing zero housing units and zero people, obtain the geography codes from the 1990 Geographic Reference File.

1990 Urban/Rural
1990 Urban Area Size Code

Append these codes to every record corresponding to the 1990 tabulation block plus suffix in the Merged Data File. See Attachment D for a list of the possible values of the Urban Area Size Code.

2. Assign 1990 Tabulation Block Information to 2000 Collection Blocks

The Merged Data File now has information about the 1990 tabulation blocks. This section describes the procedure used to create estimates for the 2000 collection blocks. An example displayed in Tables 1 and 2 illustrates this process.

a. At this point the Merged Data File should still be sorted by county, 1990 census tract, 1990 tabulation block and suffix. For each 1990 tabulation block plus suffix, sum the number of HUs in all 2000 collection blocks associated with that tabulation block. Use the 'Number of HUs in Block' field that was taken from the Block Cluster File. Append this total to each record where the tabulation block appears in the Merged Data File.

A 2000 collection block would be considered associated with a 1990 tabulation block if a portion of the collection block falls within the boundaries of the tabulation block.

b. For each record in the Merged Data File do the following:

i. Create a proportion, P, defined to be the number of HUs in the 2000 collection block divided by the total HUs associated with the 1990 tabulation block. Round P to 6 decimal places (0.0000005 rounds up to 0.000001). If the Total 2000 HUs associated with a 1990 tabulation block is equal to zero then P is undefined. In these cases set P = 0.

$$P = \frac{2000\ collection\ block\ HUs}{Total\ 2000\ HUs\ associated\ with\ 1990\ tabulation\ block}$$

ii. Multiply this proportion (P) by the number of people in each of the 12 demographic/tenure groups, the total number of owners and renters, the number of HUs, and the number of occupied HUs in the 1990 tabulation block.

iii. Append these 16 counts (the products computed in ii.) to the Merged Data File. Round these figures to four decimal places (0.00005 rounds up to 0.0001).

c. Sort the Merged Data File by county and 2000 collection block.

d. For each 2000 collection block, sum each demographic/tenure group, the estimated numbers of owners and renters, and the estimated number of HUs and occupied HUs across all records for that 2000 collection block. Round to the nearest integer (0.5 rounds to 1). These summations are the estimated population counts for the 2000 collection blocks. Append these 16 estimates to each record for the 2000 collection block in the Merged Data File.

Now the Merged Data File contains two HU counts for each 2000 collection block. One is based on the HU count from block clustering and will be referred to as the HU count. The other is based on the 1990 tabulation block information and will be referred to as the estimated HU count. This second count will be used later by the Long Form Sample Design and Estimation Team.

e. Calculate the total number of estimated people in each 2000 collection block by summing the estimated number of owners and the estimated number of renters. Append this value to each record for the 2000 collection block in the Merged Data File.

The following example depicts a site containing five 1990 tabulation blocks and five 2000 collection blocks. The example is limited to two demographic/tenure groups for ease of demonstration.

Table 1 shows how to assign population characteristics to 2000 collection block parts (steps 2a and 2b). Table 2 shows how to sum across these 2000 collection block parts to get 2000 collection block totals (step 2d).

Table 1. Assigning Population Characteristics to 2000 Collection Block Parts

1990 Tabulation Block	2000 Collection Block	Housing units in Collection block	1990 Tabulation Block Persons		Total housing units in all Collection Blocks associated with Tabulation Block	P	Persons assigned to 2000 collection block	
			Black Owner	Asian Renter			Black Owner	Asian Renter
117	14157	36	60	100	48	36/48	45	75
117	14158	12				12/48	15	25
118	14157	36	144	0	72	36/72	72	0
118	14158	12				12/72	24	0
118	14167	24				24/72	48	0
119	14157	36	100	50	36	36/36	100	50
120	14159	48	5	30	48	48/48	5	30
121	14158	12	120	12	84	12/84	17.1429	1.7143
121	14163	72				72/84	102.8571	10.2857

Table 2. Summing of 2000 Collection Block Parts

2000 Collection Block		1990 Tabulation Block					
		117	118	119	120	121	Total
14157	Black Owners	45	72	100	0	0	217
	Asian Renters	75	0	50	0	0	125
14158	Black Owners	15	24	0	0	17.1429	56
	Asian Renters	25	0	0	0	1.7143	27
14159	Black Owners	0	0	0	5	0	5
	Asian Renters	0	0	0	30	0	30
14163	Black Owners	0	0	0	0	102.8571	103
	Asian Renters	0	0	0	0	10.2857	10
14167	Black Owners	0	48	0	0	0	48
	Asian Renters	0	0	0	0	0	0

3. Assign 1990 Geography to 2000 Collection Blocks

The Merged Data File has the 1990 Urban/Rural and 1990 Urban Area Size variables for each 1990 tabulation block. This section describes how to assign these 1990 geographical variables to the 2000 collection blocks.

a. Sort the blocks within the Merged Data File by county, 1990 tract, and 1990 tabulation block number and suffix.

b. For each 1990 tabulation block record on the Merged Data File create a new variable, 1990 tabulation block urbanicity, using the variable 1990 Urban Area Size Code (see Attachment D for possible values of this variable) and 1990 Urban/Rural. This new variable will have three values:

1 = Urban Area with 1990 population ≥ 250,000
2 = Other Urban Area
3 = Non-Urban Area

Use the following algorithm to create 1990 tabulation block urbanicity:

If 1990 Urban Area Size ≥ 19 then
    1990 tabulation block urbanicity = 1
else if 1990 urban/rural = 0 then
    1990 tabulation block urbanicity = 2
else 1990 tabulation block urbanicity = 3

c. Sort the Merged Data File by county and 2000 collection block.

d. For each 2000 collection block, compare the number of people in all the 1990 tabulation blocks associated with the 2000 collection block. Choose the 1990 tabulation block with the most people. Assign that tabulation block's urbanicity to the estimated urbanicity of the 2000 collection block. Make this change for all the 1990 tabulation block/2000 collection block associations that involve the 2000 collection block.

If there is a tie for having the most people, then compare the 1990 tabulation block urbanicity values of all blocks involved in the tie. Assign the lowest value (i.e. closest to 0) to the estimated 1990 urbanicity of the 2000 collection block. Again, make this change for all 1990 tabulation block/2000 collection block associations that involve the 2000 collection block.

4. Produce Merged Data File Verification Outputs

Provide the DSSD access to the Merged Data Files for verification. The DSSD will verify the creation of the Merged Data File for eight selected counties. See Attachment E for a list of counties. See Attachment F for the layout of the Merged Data File.

B. Create the Universe File

At this point we have all the data we need on the Merged Data File and we will begin using that information to create the Universe File -- a block level file with block and cluster information. The basis of this file will be the Block Cluster File (see Attachment B for a layout). The following are the steps to create the Universe File which include assigning stratification and sampling variables.

1. Calculate Cluster Level Variables

The Merged Data File has 2000 collection block variables. These variables will be transferred to the Block Cluster File and then summarized to the cluster level.

a. Sort the Merged Data File by county and 2000 collection block.

b. Sort the Block Cluster File by county and 2000 collection block. From this point on, the Block Cluster File will be referred to as the Universe File.

c. For each 2000 collection block, append the estimated number of people in each of the 12 demographic/tenure groups, the two tenure groups, and the estimated number of HUs and occupied HUs to the 2000 collection block record in the Universe File. These counts can be found in the Merged Data File.

   The Universe File now contains two HU counts for each 2000 collection block. One was used to do the block clustering and will be referred to as the HU count. The other is based on the 1990 tabulation block information and will be referred to as the estimated HU count. This second count will be used later by the Long Form Sample Design and Estimation Team.

d. Append the estimated 1990 urbanicity of each 2000 collection block to the Universe File. This information can be found on the Merged Data File.

e. Sort the Universe File by county, block cluster number, and 2000 collection block.

f. Sum each of the 12 demographic/tenure counts, the two tenure counts, and the estimated HU and occupied HU counts across all 2000 collection blocks in a 2000 collection block cluster. On the Universe File, append these 16 block cluster totals to each collection block in the block cluster.

g. For each 2000 collection block cluster, examine the current HU count (from block clustering) for all collection blocks in the cluster. Choose the estimated 1990 urbanicity of the 2000 collection block with the most HUs. Assign that value to the estimated urbanicity of the 2000 collection block cluster for each collection block in the cluster.

If two or more blocks tie for the most HUs then examine the estimated 1990 urbanicity of all those that tie. Choose the lowest value (closest to 0) among them. Assign that value to the estimated 1990 urbanicity of the 2000 collection block cluster for each collection block in the cluster.

2. Create Sampling and Sorting Variables

Now we will create some variables that will be used during the sampling and sorting operations. Assign the cluster variable created below to all 2000 collection blocks in the cluster.

a. Since different sampling operations will be performed on different size block clusters, we want to classify clusters into three size categories -- small, medium, and large. Create a variable for cluster size as follows.

Table 3. Block Cluster Size Classification Rules

IF	THEN
HU count (From block clustering)	Size Category
0-2	1
3-79	2
80+	3

b. For 24 states, DC, and Puerto Rico there will be three sampling strata. These are the areas with few or no American Indians on reservation. For the other 26 states that have an American Indian population on reservation there will be four sampling strata. Use the number of sampling strata variable on the Sample Size Input File (ACE2000_TB.FIN). See Attachment K for a file layout. Attachment G also has a listing of which states fall into which category.

For each 2000 collection block, create a variable which record the number of sampling strata in the state. In states with three strata (those without American Indian Reservation sample), the strata will be defined as follows:

Table 4. Sampling Strata: States without American Indian Reservation Sample

IF	THEN
Size Category	Sampling stratum
1	1
2	2
3	3

In states with four strata (those with AIR sample), small block clusters on Reservations will be grouped with other small block clusters. Medium and large block clusters on Reservations will be placed in the same strata. Documentation for the AIR sample is forthcoming.

Table 5. Sampling Strata: States with American Indian Reservation Sample

IF		THEN
Size Category	American Indian Country Indicator	Sampling stratum
1	0, 1, or 2	1
2	0 or 2	2
2	1	4
3	0 or 2	3
3	1	4

c. To help make sure we have a balanced representation of all groups in our sample we are going to create a sort variable based on the demographic/tenure make-up of each block cluster. Create this demographic/tenure variable and a label as follows.

i. For each 2000 block cluster, divide the number of people in each of the 12 demographic/tenure groups by the total number of people to determine the population proportion of each group. Use the same method to also determine the population proportion of renters and owners in the block

cluster. Round each value to three decimal places (x.0005 rounds to x.001). Append these values to the Universe File.

ii. For the 50 states and DC, assign the cluster to a demographic/tenure group based on the following ordered rules. These rules are based on research that will be documented at a later date.

The first rule that a cluster satisfies is the group to which the cluster is assigned. Once a cluster is assigned, stop processing that cluster.

Table 6. Demographic/Tenure Group Assignment Rules for the 50 States and the District of Columbia

	IF	THEN	
Order	Criteria	Dem/Ten Group Code	Dem/Ten Group Label
1	Proportion of Hawaiian and Pacific Islander Renters ≥ 0.10	1	PR
2	Proportion of Hawaiian and Pacific Islander Owners ≥ 0.10	2	PO
3	Proportion of American Indian and Alaska Native Renters ≥ 0.10	5*	IR
4	Proportion of American Indian and Alaska Native Owners ≥ 0.10	6*	IO
5	Proportion of Asian Renters ≥ 0.20	3*	AR
6	Proportion of Asian Owners ≥ 0.20	4*	AO
7	Proportion of Hispanic Renters ≥ 0.20	7	HR
8	Proportion of Hispanic Owners ≥ 0.20	8	HO
9	Proportion of Black Renters ≥ 0.25	9	BR
10	Proportion of Black Owners ≥ 0.25	10	BO
11	Proportion of Other Renters ≥ 0.30	11	OR
12	all else	12	OO

*Smaller demographic groups are given precedence in the assignment process. Within each race group, renter is also given precedence over owner. This order does not coincide with the demographic/tenure group codes, which are assigned so that certain groups will be sorted together before selection.

iii. For Puerto Rico, assign the cluster to a demographic/tenure group based on the following rules. The first rule that a cluster satisfies is the group the cluster is assigned to. [Note: The groups are listed in ascending order based on the number of members.]

Table 7. Demographic/Tenure Group Assignment Rules for Puerto Rico

	IF	THEN	
Order	Criteria	Dem/Ten Group Code	Dem/Ten Group Label
1	Proportion of Renters ≥ 0.45	13	TR
2	all else	14	TO

d. Assign each collection block a region code and a division code. This information is contained in the Sample Size Input File and in Attachment G. These are based on the standard definitions used by the GEO.

3. Produce Universe File Verification Outputs

a. We will evaluate how the Demographic/Tenure Group variable classifies the block clusters. As such, for each state, DC and Puerto Rico, tally the demographic/tenure counts and the HU and block cluster counts in a sampling strata for each Demographic/Tenure Group compute several person and HU counts. Write this information to a file. See Attachment H for a file layout and example.

b. We would also like to evaluate the estimated 1990 urbanicity variable. For each state, calculate the proportion of 2000 collection blocks in each estimated 1990 urbanicity group. Then calculate the number of people that live in blocks with each of the three classifications. Also, calculate the proportion of 2000 collection block clusters that fall into each urbanicity group and the number of people that live in block clusters which fall into each urbanicity group. Write this information to a file. See Attachment I for a layout.

c. Provide the DSSD access to the Universe File for verification. See Attachment J for a file layout of the Universe File.

IV. Creation of the Block Cluster Sampling Parameter File

The parameters for selecting the initial sample of block clusters are specified in this section. These parameters will be an input to a forthcoming memorandum specifying the selection of the block cluster sample.

The block cluster sampling will be done separately for each of the four sampling strata within each state. As part of the Universe File creation in Section III, all clusters were assigned to one of four sampling strata: 1) Small block clusters, 2) Medium block clusters, 3) Large block clusters, and 4) AIR block clusters[2]. Each of these four categories will be sampled at different rates during the initial block cluster sampling.

Approximately 25,000 block clusters were originally allocated to the 50 states and the District of Columbia[3]. This allocation was recently updated, resulting in the target number of medium and large sample block clusters for each state as shown in Attachment G. Attachment G also contains the target number of small block clusters for each state, and the target number of clusters for Puerto Rico. These numbers, which are needed for the calculations that follow, can be obtained from the Sample Size Input File, which is described in the Input Section below. This file will be the basis for the Block Cluster Sampling Parameter File. The procedures for calculating the first-step sampling parameters for each state are below. Do these steps separately for each sampling stratum.

A. Calculation of the Sampling Parameters

The calculation of the sampling parameters requires two basic pieces of information for each state and sampling stratum: the sample size and the universe count. In general, the take-every is the ratio of the universe count to the sample size. The calculation of the large block sampling stratum take-every has one additional calculation which controls the sample so that the expected listing does not exceed the budgeted listing.

1. Sample Sizes

Obtain the following information for the state from the Sample Size Input File (see Attachment K for a layout):

---

[2]Not all states will have block clusters classified into sampling stratum 4. See Attachment G.

[3]Schindler, E. (1998), "Allocation of the ICM Sample to the States for Census 2000," Proceedings of the Section on Survey Research, American Statistical Association.

Budgeted Listing for medium and large clusters, BLIST
Target number of clusters for small sampling stratum, TCLUSTS
Target number of clusters for the medium and large sampling strata, TCLUST
Target number of clusters in AIR stratum, TCLUSTA

2. Universe File Counts

For each state use the Universe File to produce the total number of block clusters and the total number of HUs for each of the four sampling strata. The counts are:

Total number of housing units in small clusters, NHUS
Total number of small block clusters, NCLUSTS
Total number of housing units in medium clusters, NHUM
Total number of medium clusters, NCLUSTM
Total number of housing units in large clusters, NHUL
Total number of large clusters, NCLUSTL
Total number of housing units in AIR, NHUA
Total number block clusters in AIR, NCLUSTA.

3. Allocate Block Clusters to Sampling Strata

For the small block cluster and AIR sampling strata, the target clusters are given on the Sample Size Input File. For the medium and large sampling strata, a combined target has been provided on the Sample Size Input File. This combined target needs to be proportionally allocated to the large and medium sampling strata.

Calculate the target number of medium block clusters and the target number of large block clusters by proportionally allocating the combined target to the medium and large sampling strata based on HUs.

$$\text{TCLUSTM} = \frac{\text{NHUM}}{\text{NHUM} + \text{NHUL}} \times \text{TCLUST}$$

$$\text{TCLUSTL} = \frac{\text{NHUL}}{\text{NHUM} + \text{NHUL}} \times \text{TCLUST}$$

4. Check if Listing is Within Budget

A listing adjustment is applied to the large block cluster take-every to aid in controlling the listing workload. If the preliminary expected number of

HUs is larger than the budgeted listing, an adjustment needs to be made to the large block sampling rate to reduce the expected number of HUs.

PRELIST is the preliminary number of expected HUs to be listed from the medium and large sampling strata.

$$\text{PRELIST} = \text{TCLUSTM} \times \frac{\text{NHUM}}{\text{NCLUSTM}} + \text{TCLUSTL} \times \frac{\text{NHUL}}{\text{NCLUSTL}}$$

Compare the preliminary expected listing to the budgeted listing and calculate the listing adjustment, LISTADJ. Round LISTADJ to 6 decimal places.

If BLIST < PRELIST, then

$$\text{LISTADJ} = \frac{\text{BLIST} - \text{TCLUSTM} \times \frac{\text{NHUM}}{\text{NCLUSTM}}}{\text{TCLUSTL} \times \frac{\text{NHUL}}{\text{NCLUSTL}}}$$

If BLIST ≥ PRELIST, then LISTADJ = 1.000000

5. Calculate Sampling Parameters

Calculate the take-everys and random starts for the four sampling strata. The take-everys are calculated by dividing the total number of clusters by the target number of clusters. Then if necessary, the large block cluster take-every is then adjusted to comply with the budgeted listing. Random starts are calculated by multiplying the take-every by a random number between zero and one ($0 < RN \leq 1$). For each state and sampling stratum (SS=1, ..., 4) obtain the appropriate number of clusters from step two and the corresponding target from step three. Generate a new random number for each of the four sampling strata.

Calculate the take-every:

For sampling strata 1, 2 and 4:

$$TE1_{SS} = \frac{NCLUST_{SS}}{TCLUST_{SS}}$$ (Round to 6 decimal places.)

For sampling stratum 3:

$$TE1_{SS} = \frac{NCLUST_{SS}}{TCLUST_{SS} \times LISTADJ} \quad \text{(Round to 6 decimal places.)}$$

Calculate a random start for each of the four sampling strata:

$$RS1_{SS} = RN \times TE1_{SS}. \quad \text{(Round to 6 decimal places.)}$$

B. Calculation of Expected Number of Housing Units and Clusters for Listing

These expected values will be placed on the Sample Summary File and compared to the sampling results. For each state and sampling strata obtain the total number of clusters from step two and the take-every from step five in section IV.A. Calculate the expected number of clusters and HUs for each sampling stratum:

$$ECLUSTL_{SS} = \frac{NCLUST_{SS}}{TE1_{SS}} \quad \text{(Round to the nearest integer.)}$$

$$EXPHUL_{SS} = \frac{NHU_{SS}}{TE1_{SS}} \quad \text{(Round to the nearest integer.)}$$

C. Creation of the Sample Summary File and Block Cluster Sampling Parameter File

1. Sample Summary File

The Sample Summary File is a state level file and is one of the two files used to provide a history of the ACE sample[4]. A description and layout of the file was documented in a previous memorandum, but the initial version will be created in this section. Future versions of the file will be created as updates are made following each sampling operation. Using the information calculated for each state and sampling stratum in section IV.A, create version one of the Sample Summary File, ACE2000_SSFV1.<mmddyy>. The extension <mmddyy> is the date the file is created (i.e. 123199 is the extension for a file created on December 31, 1999). Include the following variables:

1. Census Region, REGION
2. Census Division, DIV

[4]The second file used in tracking the ACE sample is the Sample Design File will be created following the initial block cluster sampling. Descriptions of these files are documented in the memorandum from Kostanich to Stoudt, "Accuracy and Coverage Evaluation (ACE) Survey: Documentation for the Sample Summary File and Sample Design File," March 1999.

3. State code (01-72 = FIPS State Code), STATE
4. Budgeted Listing for medium and large clusters, BLIST
5. Target number of medium and large block clusters, TCLUST
6. Target number of small block clusters, TCLUSTS
7. Target number of AIR block clusters, TCLUSTA
8. Total number of block clusters, NCLUST

$$NCLUST = \sum_{ss=1}^{4} NCLUST_{ss}$$

9. Total number of housing units, NHU

$$NHU = \sum_{ss=1}^{4} NHU_{ss}$$

10. Expected clusters in sample to list, ECLUSTL

$$ECLUSTL = \sum_{ss=1}^{4} ECLUSTL_{ss}$$

11. Expected housing units in sample to list, EXPHUL

$$EXPHUL = \sum_{ss=1}^{4} EXPHUL_{ss}$$

2. The Block Cluster Sampling Parameter File

The Block Cluster Sampling Parameter File contains one record for each sampling stratum in each of the 50 states, the District of Columbia, and Puerto Rico. During production, a separate file will be created for each state, the District of Columbia and Puerto Rico. These files will serve as inputs to the block cluster sample selection specification. Following the completion of the sample selection for all states, the state files will be concatenated into one file. A description of this file, ACE2000_PARABC.<mmddyy>, is provided in section V. Using the information from section IV.A for each state and sampling stratum, create the Block Cluster Sampling Parameter Files. Include the following variables:

1. Census Region, REGION
2. Census Division, DIV
3. State code (01-72 = FIPS State Code), STATE
4. Sampling Stratum, SS
5. Target number of block clusters, TCLUST
6. Total number of block clusters, NCLUST
7. Total number of housing units, NHU
8. First-step take-every, TE1
9. First-step random start, RS1

D. Block Cluster Sampling Parameter File Verification Output

Using the Block Cluster Sampling Parameter File and a listing of the random numbers generated to calculate the random starts, the DSSD will conduct an independent verification of the parameter calculations.

V. Input

The following files will be needed for producing both the Universe File and the Block Cluster Sampling Parameter File.

A. Block Cluster Files

These are block level files which contain information about which blocks are clustered together. There will be one file for each state. These files are created by the GEO as specified in the memorandum from Hogan to Marx, "Census 2000 Specifications for Block Cluster Formation," February 16, 1999. See Attachment B for a file layout.

B. 1990 Hundred Percent Edited Detail File

This file contains data collected from the 1990 census. We will get both person information and 1990 tabulation block information from this file. This is the source of person-level demographic and tenure information.

C. Block Equivalency Files

These files show the relationships between the 1990 tabulation blocks and the 2000 collection blocks. In these files there is a record for each 1990 tabulation block/2000 collection block relationship. There is one file per county. See Attachment A for a file layout.

D. 1990 Geographic Reference File

This file contains geographic and legal information about 1990 tabulation and collection blocks.

E. Sample Size Input File: ACE2000_TB.FIN

This file contains information about each state that is needed for the sampling process. This file will be provided by the DSSD. There is one record for each state, the District of Columbia, and Puerto Rico. See Attachment K for a file layout.

VI. Output

The following files will be outputs from the system.

A. Merged Data Files

This is the Block Equivalency File with additional information appended. For a layout, see Attachment F. There will be one file for each state.

B. Universe Files

This is the Block Cluster File with additional information appended. For a layout, see Attachment J. This file is at the 2000 collection block level and contains information about both the collection blocks and the collection block clusters. There will be one file per state.

C. Demographic/Tenure Group Evaluation File

This is a file which contains demographic/tenure totals and HU counts for each state and demographic/tenure group.

D. Urbanicity Verification File

This is a file which contains population totals, proportions of 2000 collection blocks, and proportions of 2000 block clusters for each of the three urbanicity categories.

E. Version One of the Sample Summary File: ACE2000_SSFV1.FIN

This is a file containing cluster and HU totals at the state level. The file has one record for each state, the District of Columbia and Puerto Rico. For a complete description and layout of the file, see the memorandum from Kostanich to Stoudt, "Accuracy and Coverage Evaluation Survey: Documentation for the Sample Summary File and Sample Design File," March 1999.

F. Block Cluster Sampling Parameter File: ACE2000_PARABC.FIN

This is a file containing block cluster sampling parameters. It will be an input to the block cluster sampling specification. The file contains one record per sampling stratum for each state, the District of Columbia and Puerto Rico. See Attachment L for a layout.

G. List of random numbers

The list contains the random numbers generated to calculate the random starts for each sampling strata. Depending on the number of sampling strata there are three or four random numbers for each state, the District of Columbia, and Puerto Rico.

cc: DSSD Census 2000 Memorandum Series Distribution List
ACE Implementation Team/Statistical Design Team Leaders List
Sample Design Team
Sam Hawala (PRED)
Andrew Zbikowski (PRED)

Attachment A

Block Equivalency File Layout

Variable Description	Name	Places
1990 Tabulation State	STATE90	1-2
1990 Tabulation County	COUNTY90	4-6
1990 Tabulation Census Tract	TRACT90	8-11
1990 Tabulation Tract Suffix	TRASUF90	12-14
1990 Tabulation Block	TBLOCK90	15-17
1990 Tabulation Suffix	TBSUF90	18-18
2000 Collection State	STATE2K	20-21
2000 Collection County	COUNTY2K	23-25
2000 Collection Block	CBLOCK2K	27-31

Block Cluster File Layout

Variable Description	Name	Places
State	STATE	1-2
County	COUNTY	3-5
Interim Tract (a.k.a. pseudo-tract)	ITRACT	6-11
Block Number	COLBLOCK	12-16
Blank		17-17
Cluster Number (geography not ACE)	GCLUS	18-22
Blank		23-23
Cluster Size code	CLUSSIZE	24-24
1 = Clusters with 0 HUs		
2 = Clusters with 1 HUs		
3 = Clusters with 2 HUs		
4 = Clusters with between 3 and 5 HUs		
5 = Clusters with between 6 and 9 HUs		
6 = Clusters with between 10 and 19 HUs		
7 = Clusters with between 20 and 29 HUs		
8 = Clusters with between 30 and 79 HUs		
9 = Clusters with 80 or more HUs		
Blank		25-25
Block Area (Sq. Miles)	BAREA	26-33
Blank		34-34
Block Perimeter (Miles)	BPERIM	35-40
Blank		41-41
Block Cluster Area (Sq. Miles)	BCAREA	42-49
Blank		50-50
Block Cluster Perimeter (Miles)	BCPERIM	51-56
Number of HUs in cluster	NHU	57-61
Number of HUs in block	NHUBLOCK	62-66
Block TEA	TEA	67-67
1 = Mailout/Mailback		
2 = Update/Leave		
3 = List/Enumerate		
5 = Rural Update/Enumerate		
6 = Military		
7 = Urban Update/Leave		
8 = Update/Leave to Mailout/Mailback conversions		
9 = Mailout/Mailback to Update/Leave conversions		

TEA Group for Block Cluster A= Mailout/Mailback or Urban Update/Leave or Update/Leave to Mailout/Mailback conversions B= Update/Leave or Rural Update/Enumerate C=List/Enumerate D=Military E=Mailout/Mailback to Update/Leave conversions	TEABC	68-68
2000 MAF HUs count ' ' Blank if no HU count available	NHUM	69-73
1990 ACF HUs count ' ' Blank if no HUs count available	NHU90	74-78
Housing Unit Count Indicator 1 = from 2000 MAF 2 = from 1990 ACF	HUIND	79-79
Invisible Boundary Collapse Indicator 0 = No 1 = Yes (Collapsing across Invisible Boundary in BC)	INV	80-80
American Indian Country Indicator 0 = No American Indian Country 1 = American Indian Reservation/trust land 2 = Tribal jurisdiction statistical area/ Alaska Native Village statistical area/ tribal designated statistical area	AICIND	81-81
Military Indicator 0 = No Military Area 1 = Block contains Military Area	MILIND	82-82
Collapsed Enclosed Block Indicator 0 = Otherwise 1 = An enclosed block has been forced to collapse	CEBI	83-83

Definitions of the Demographic and Tenure Characteristics

Using variables Q4 (race), Q7 (Hispanic origin), and H4 (tenure) from the 1990 HEDF we can assign any person to one of the 12 demographic/tenure groups.

Questions Q4 and Q7 are used to determine each person's demographic group. Question Q4 asks for the respondent's race. Write-in responses are coded to the values 001 to 946. Values 971 to 986 are used for marked responses which include white, black, American Indian, Eskimo, Aleut, Chinese, etc. Question Q7 asks for a person's Hispanic origin. Values 2-5 indicate different Hispanic origins (Cuban, Puerto Rican, etc.)

The six demographic categories have been designed to be mutually exclusive. So, if a person is of Hispanic Origin that person would be classified as Hispanic regardless of race. The following algorithm has been developed to classify each person.

American Indian and Alaska Native =	([Q4 ≥ 000 AND Q4 ≤ 599] OR [Q4 ≥ 935 AND Q4 ≤ 940] OR [Q4 ≥ 941 AND Q4 ≤ 970] OR [Q4 ≥ 973 AND Q4 ≤ 975]) AND Q7≤ 1
Asian =	([Q4 ≥ 600 AND Q4 ≤ 652] OR [Q4 ≥ 976 AND Q4 ≤ 977] OR [Q4 ≥ 979 AND Q4 ≤ 982] OR Q4 = 985) AND Q7≤ 1
Black =	([Q4 ≥ 870 AND Q4 ≤ 934] OR Q4 = 972) AND Q7≤ 1
Hawaiian and Pacific Islander =	([Q4 ≥ 653 AND Q4 ≤ 699] OR Q4 = 978 OR Q4 = 983 OR Q4 = 984) AND Q7≤ 1
Hispanic =	Q7 ≥ 2
White and Other =	Remaining records (any records left over after all other race/ethnic group criteria are exhausted)

Tenure is defined by the H4 variable as follows:

Owner = H4 = 1 OR H4 = 2

Renter = H4 = 3 OR H4 = 4

To determine the number of HUs and the number of occupied HUs use the variable HC1. The variable HC1 has the following values:

0 = occupied
1 = for rent
2 = for sale only
3 = rented or sold, not occupied
4 = for seasonal/recreational/occasional use
5 = for migratory workers
6 = other vacant

Occupied = HC1 = 0
Vacant = HC1 ≠ 0

Attachment D

Size of Urban Area

The following are the possible values for the urban area size of a 1990 tabulation block.

Code	Number of People
0	not in universe
1	0
2	1 - 24
3	25 - 99
4	100 - 199
5	200 - 249
6	250 - 299
7	300 - 499
8	500 - 999
9	1,000 - 1,499
10	1,500 - 1,999
11	2,000 - 2,499
12	2,500 - 4,999
13	5,000 - 9,999
14	10,000 - 19,999
15	20,000 - 24,999
16	25,000 - 49,999
17	50,000 - 99,999
18	100,000 - 249,000
19	250,000 - 499,999
20	500,000 - 999,999
21	1,000,000 - 2,499,999
22	2,500,000 - 4,999,999
23	5,000,000 or more

Attachment E

Merged Data File: Verification Counties

For verification of the Merged Data File, provide the DSSD with the files for counties listed below. These are all from wave one states.

Bear Lake County, Idaho (FIPS = 16007)
Grand Forks County, North Dakota (FIPS = 38035)
Anchorage Borough, Alaska (FIPS = 02020)
Rosebud County, Montana (FIPS = 30087)
Blue Earth County, Minnesota (FIPS = 27013)
District of Columbia (the entire district is considered a county) (FIPS = 11001)
Shannon County, South Dakota (FIPS = 46113)
Menominee County, Wisconsin (FIPS = 55078)

Merged Data File Layout

Variable Description	Name	Places
1990 Tabulation State	STATE90	1-2
1990 Tabulation County	COUNTY90	4-6
1990 Tabulation Census Tract	TRACT90	8-11
1990 Tabulation Tract Suffix	TRASUF90	12-14
1990 Tabulation Block	TBLOCK90	15-17
1990 Tabulation Suffix	TBSUF90	18-18
2000 Collection State	STATE2K	20-21
2000 Collection County	COUNTY2K	23-25
2000 Collection Block	CBLOCK2K	27-31
Blank		32-35
2000 Collection Block Housing Unit Count	CB2KHU	36-40
Geography Block Cluster Number	GCLUST	41-45
Blank		46-49
1990 Tabulation Block Number of:		
Hawaiian and Pacific Islander Renter	PIR90TAB	50-54
Hawaiian and Pacific Islander Owner	PIO90TAB	55-59
American Indian and Alaska Native Renter	IR90TAB	60-64
American Indian and Alaska Native Owner	IO90TAB	65-69
Asian Renter	AR90TAB	70-74
Asian Owner	AO90TAB	75-79
Hispanic Renter	HR90TAB	80-84
Hispanic Owner	HO90TAB	85-89
Black Renter	BR90TAB	90-94
Black Owner	BO90TAB	95-100
White and Other Renter	OR90TAB	100-104
White and Other Owner	OO90TAB	105-109
Total Renters	R90TAB	110-114
Total Owners	O90TAB	115-119
Total Housing Units	HU90TAB	120-124
Occupied Housing Units	OHU90TAB	125-129
Total People (Non-GQ)	POP90TAB	130-134
Blank		135-145
1990 Urban/Rural (UR)	UR90	146-146
1990 Urban Area Size (UASZ)	UASZ90	147-148
# of 2000 collection block HUs associated with 1990 tabulation block	HU2KA90	149-153
Proportion of HUs in 2000 collection block associated with 1990 tabulation block	P	154-161

Blank		162-170
1990 Tab/2000 Collection Association Estimated Number of:		
Hawaiian and Pacific Islander Renter	PIRASSOC	171-175
Hawaiian and Pacific Islander Owner	PIOASSOC	176-180
American Indian and Alaska Native Renter	IRASSOC	181-185
American Indian and Alaska Native Owner	IOASSOC	186-190
Asian Renter	ARASSOC	191-195
Asian Owner	AOASSOC	196-200
Hispanic Renter	HRASSOC	201-205
Hispanic Owner	HOASSOC	206-210
Black Renter	BRASSOC	211-215
Black Owner	BOASSOC	216-220
White and Other Renter	ORASSOC	221-225
White and Other Owner	OOASSOC	226-230
Total Renters	RASSOC	231-235
Total Owners	OASSOC	236-240
Total Housing Units	HUASSOC	241-245
Occupied Housing Units	OHUASSOC	246-250
Blank		251-260
2000 Collection Block Estimated Number of:		
Hawaiian and Pacific Islander Renter	ECOLPIR	261-265
Hawaiian and Pacific Islander Owner	ECOLPIO	266-270
American Indian and Alaska Native Renter	ECOLIR	271-275
American Indian and Alaska Native Owner	ECOLIO	276-280
Asian Renter	ECOLAR	281-285
Asian Owner	ECOLAO	286-290
Hispanic Renter	ECOLHR	291-295
Hispanic Owner	ECOLHO	296-300
Black Renter	ECOLBR	301-305
Black Owner	ECOLBO	306-310
White and Other Renter	ECOLOR	311-315
White and Other Owner	ECOLOO	316-320
Total Renters	ECOLR	321-325
Total Owners	ECOLO	326-330
Total Housing Units	ECOLHU	331-335
Occupied Housing Units	ECOLOHU	336-340
Blank		341-350
Estimated total population of 2000 collection block	ECOLPOP	351-355
1990 tabulation block urbanicity	90TABURB	356-356
Estimated 1990 urbanicity of 2000 collection block	ECOLURB	357-357

Attachment G

Sampling Information for Each State

Region	Division	State	Number of Strata	Budgeted Listing	AIR Cluster Target	Small Cluster Target	Medium and Large Cluster Target	First Block Cluster Number
3	6	Alabama*	3	25,347	0	116	417	41001
4	9	Alaska	4	27,196	1	20	334	91001
4	8	Arizona	4	48,451	110	86	492	81001
3	7	Arkansas	3	24,744	0	90	494	51001
4	9	California	4	284,076	14	184	2,753	92001
4	8	Colorado	4	37,965	2	83	479	82001
1	1	Connecticut*	3	30,039	0	20	377	11001
3	5	Delaware	3	21,610	0	20	413	31001
3	5	DC	3	53,369	0	20	384	32001
3	5	Florida	4	62,845	1	145	520	33001
3	5	Georgia*	3	37,384	0	154	399	34001
4	9	Hawaii	3	45,059	0	20	300	97001
4	8	Idaho	4	19,157	6	54	412	83001
2	3	Illinois	3	31,571	0	185	430	61001
2	3	Indiana	3	15,925	0	140	275	62001
2	4	Iowa*	3	14,108	0	147	300	71001
2	4	Kansas	4	16,281	1	193	300	72001
3	6	Kentucky	3	29,621	0	96	447	42001
3	7	Louisiana*	3	37,378	0	65	595	52001
1	1	Maine	4	16,572	1	38	309	12001
3	5	Maryland	3	41,107	0	36	368	35001
1	1	Massachusetts*	3	27,255	0	38	375	13001
2	3	Michigan	4	24,128	5	122	379	63001
2	4	Minnesota	4	19,091	10	141	300	73001
3	6	Mississippi	4	19,990	3	81	402	43001
2	4	Missouri	3	19,807	0	162	300	74001
4	8	Montana	4	17,969	24	67	420	84001
2	4	Nebraska	4	14,177	3	142	300	75001
4	8	Nevada	4	63,031	5	46	468	85001
1	1	New Hampshire	3	21,128	0	25	307	14001
1	2	New Jersey**	3	37,394	0	39	461	21001
4	8	New Mexico	4	32,242	70	108	481	86001
1	2	New York	4	143,949	5	143	1,261	22001
3	5	North Carolina	4	26,717	4	143	400	36001
2	4	North Dakota	4	15,738	12	121	300	76001
2	3	Ohio	3	30,790	0	132	421	64001
3	7	Oklahoma	4	25,328	8	142	426	53001
4	9	Oregon	4	20,577	3	86	320	98001

Region	Division	State	Number of Strata	Budgeted Listing	AIR Cluster Target	Small Cluster Target	Medium and Large Cluster Target	First Block Cluster Number
1	2	Pennsylvania	3	34,920	0	180	585	25001
1	1	Rhode Island*	3	23,557	0	20	373	15001
3	5	South Carolina*	3	26,709	0	95	422	37001
2	4	South Dakota	4	14,227	27	106	300	77001
3	6	Tennessee	3	30,255	0	133	433	44001
3	7	Texas	4	176,234	1	349	1,945	54001
4	8	Utah	4	32,777	7	38	478	87001
1	1	Vermont	3	17,009	0	21	300	16001
3	5	Virginia*	3	37,114	0	98	371	38001
4	9	Washington	4	26,832	17	73	332	99001
3	5	West Virginia	3	17,557	0	46	300	39001
2	3	Wisconsin	4	14,470	10	119	275	65001
4	8	Wyoming	4	18,293	5	72	418	88001
		United States		1,949,070	355	5000	24,651	
0***	0***	Puerto Rico	3	46,700	0	96	480	17001

* States contain AIR population, but not AIR sampling stratum. Air people will be given a chance of selection in the general state sample.

** New Jersey AIR reservations had no population in 1990.

*** Puerto Rico is an outlying area. Because of this it has no region or division code. Therefore, we assign the code 0.

Demographic/Tenure Group Evaluation File Layout

Variable Description	Name	Places
State	STATE	1-2
Demographic/Tenure group code (for Puerto Rico only include codes 13-14, for all others only include codes 1-12)	DTCODE	4-5
Demographic/Tenure group label	DTLABEL	7-8
For each of the following counts record the number in this demographic group in this state		
# of small block clusters	SMCLUS	10-17
# of medium block clusters	MEDCLUS	18-25
# of large block cluster	LARCLUS	26-33
# of housing units in small block clusters	HUSMC	34-41
# of housing units in medium block cluster	HUMC	42-49
# of housing units in large block cluster	HULC	50-57
Est. # of Hawaiian and Pacific Islander Renters	ESTPR	58-65
Est. # of Hawaiian and Pacific Islander Owners	ESTPO	66-73
Est. # of Asian Renters	ESTAR	74-81
Est. # of Asian Owners	ESTAO	82-89
Est. # of Am. Indian Renters	ESTIR	90-97
Est. # of Am. Indian Owners	ESTIO	98-105
Est. # of Hispanic Renters	ESTHR	106-113
Est. # of Hispanic Owners	ESTHO	114-121
Est. # of Black Renters	ESTBR	122-129
Est. # of Black Owners	ESTBO	130-137
Est. # of White and Other Renters	ESTOR	138-145
Est. # of White and Other Owners	ESTOO	146-153
Est. # of Renters	ESTR	154-161
Est. # of Owners	ESTO	162-169
Estimated total population	ESTPOP	170-178

The following is an example of how the file will look.

State	Demographic/Tenure Group Code	Demographic/Tenure Group Label	Small Clusters	. . . . # of HUs in Large Cluster	PR	PO	AR..
AL	1	PR	5	548	1,000	578	45..
AL	2	PO	7	0	456	984	74..
AL	3	AR					
AL . . .							
.							
AK	1	PR . . .					
AK	2	PO . . .					
.							
AR	1.....						

Urbanicity Verification File Layout

In order to verify the assignment of the urbanicity variable we would like a file with the following layout. If after either of the transitions, from 1990 tabulation to 2000 collection or from 2000 collection to 2000 cluster, there are great differences in the proportions then we will know that we should examine the transition further. The urbanicity values will also be verified on a micro-level in each of the selected counties.

Variable Description	Name	Places
State	STATE	1-2
Proportion of 2000 collection blocks with urbanicity = 1	PCBU1	3-7
Proportion of 2000 collection blocks with urbanicity = 2	PCUB2	8-12
Proportion of 2000 collection blocks with urbanicity = 3	PCUB3	13-17
People who live in 2000 collection blocks with urbanicity = 1	PEPCUB1	18-28
People who live in 2000 collection blocks with urbanicity = 2	PEPCUB2	29-39
People who live in 2000 collection blocks with urbanicity = 3	PEPCUB3	40-50
Blank		51-59
Proportion of 2000 collection blocks clusters with urbanicity = 1	PCLUU1	60-64
Proportion of 2000 collection blocks clusters with urbanicity = 2	PCLUU2	65-69
Proportion of 2000 collection blocks clusters with urbanicity = 3	PCLUU3	70-74
People who live in 2000 collection blocks clusters with urbanicity = 1	PEPCLUU1	75-85
People who live in 2000 collection blocks clusters with urbanicity = 2	PEPCLUU2	86-96
People who live in 2000 collection blocks clusters with urbanicity = 3	PEPCLUU3	97-107

Universe File Layout

Variable Description	Name	Places
State	STATE	1-2
County	COUNTY	3-5
Interim Tract (a.k.a. pseudo-tract)	ITRACT	6-11
Block Number	COLBLOCK	12-16
Blank		17-17
Cluster Number (geography not ACE)	GCLUS	18-22
Blank		23-23
Cluster Size code	CLUSSIZE	24-24
1 = Clusters with 0 HUs		
2 = Clusters with 1 HUs		
3 = Clusters with 2 HUs		
4 = Clusters with between 3 and 5 HUs		
5 = Clusters with between 6 and 9 HUs		
6 = Clusters with between 10 and 19 HUs		
7 = Clusters with between 20 and 29 HUs		
8 = Clusters with between 30 and 79 HUs		
9 = Clusters with 80 or more HUs		
Blank		25-25
Block Area (Sq. Miles)	BAREA	26-33
Blank		34-34
Block Perimeter (Miles)	BPERIM	35-40
Blank		41-41
Block Cluster Area (Sq. Miles)	BCAREA	42-49
Blank		50-50
Block Cluster Perimeter (Miles)	BCPERIM	51-56
Number of HUs in cluster	NHU	57-61
Number of HUs in block	NHUBLOCK	62-66
Block TEA	TEA	67-67
1 = Mailout/Mailback		
2 = Update/Leave		
3 = List/Enumerate		
5 = Rural Update/Enumerate		
6 = Military		
7 = Urban Update/Leave		
8 = Update/Leave to Mailout/Mailback conversions		
9 = Mailout/Mailback to Update/Leave conversions		

TEA Group for Block Cluster A= Mailout/Mailback or Urban Update/Leave or Update/Leave to Mailout/Mailback conversions B= Update/Leave or Rural Update/Enumerate C=List/Enumerate D=Military E=Mailout/Mailback to Update/Leave conversions	TEABC	68-68
2000 MAF HUs count ' ' Blank if no HU count available	NHUM	69-73
1990 ACF HUs count ' ' Blank if no HUs count available	NHU90	74-78
Housing Unit Count Indicator 1 = from 2000 MAF 2 = from 1990 ACF	HUIND	79-79
Invisible Boundary Collapse Indicator 0 = No 1 = Yes (Collapsing across Invisible Boundary in BC)	INV	80-80
American Indian Country Indicator 0 = No American Indian Country 1 = American Indian Reservation/trust land 2 = Tribal jurisdiction statistical area/ Alaska Native Village statistical area/ tribal designated statistical area	AICIND	81-81
Military Indicator 0 = No Military Area 1 = Block contains Military Area	MILIND	82-82
Collapsed Enclosed Block Indicator 0 = Otherwise 1 = An enclosed block has been forced to collapse	CEBI	83-83

Blank		84-90
2000 Collection Block Estimated Number of:		
Hawaiian and Pacific Islander Renter	ECOLPIR0	91-95
Hawaiian and Pacific Islander Owner	ECOLPIO	96-100
American Indian and Alaska Native Renter	ECOLIR	101-105
American Indian and Alaska Native Owner	ECOLIO	106-110
Asian Renter	ECOLAR	111-115
Asian Owner	ECOLAO	116-120
Hispanic Renter	ECOLHR	121-125
Hispanic Owner	ECOLHO	126-130
Black Renter	ECOLBR	131-135
Black Owner	ECOLBO	136-140
White and Other Renter	ECOLOR	141-145
White and Other Owner	ECOLOO	146-150
Total Renters	ECOLR	151-155
Total Owners	ECOLO	156-160
Total Housing Units	ECOLHU	161-165
Occupied Housing Units	ECOLOHU	166-170
Total People (Non-GQ)	ECOLPOP	171-175
Estimated 1990 urbanicity of the 2000 collection block	ECOLURB	176-176
1 = Urban Area with 1990 population ≥ 250,000		
2 = Other Urban Area		
3 = Non-Urban Area		
Blank		177-180
2000 Collection Block Cluster Estimated Number of:		
Hawaiian and Pacific Islander Renter	ECLUSPIR	181-185
Hawaiian and Pacific Islander Owner	ECLUSPIO	186-190
American Indian and Alaska Native Renter	ECLUSIR	191-195
American Indian and Alaska Native Owner	ECLUSIO	196-200
Asian Renter	ECLUSAR	201-205
Asian Owner	ECLUSAO	206-210
Hispanic Renter	ECLUSHR	211-215
Hispanic Owner	ECLUSHO	216-220
Black Renter	ECLUSBR	221-225
Black Owner	ECLUSBO	226-230
White and Other Renter	ECLUSOR	231-235
White and Other Owner	ECLUSOO	236-240
Total Renters	ECLUSR	241-245
Total Owners	ECLUSO	246-250
Total Housing Units	ECLUSHU	251-255
Occupied Housing Units	ECLUSOHU	256-260
Total People (Non-GQ)	ECLUSPOP	261-265
Blank		266-275

Estimated 1990 urbanicity of 2000 block cluster	ECLUSURB	276-276
1 = Urban Area with 1990 population ≥ 250,000		
2 = Other Urban Area		
3 = Non-Urban Area		
Size Category	SIZECAT	277-277
1 = Small (0-2 HUs)		
2 = Medium (3-79 HUs)		
3 = Large (80+ HUs)		
Number of sampling strata in state	NSSINST	278-278
Sample stratum	SS	279-279
1 = Small		
2 = Medium (non-AIR)		
3 = Large (non-AIR)		
4 = American Indian Reservation		
Blank		280-285
2000 Collection Block Cluster Proportion of Population that is:		
Hawaiian and Pacific Islander Renter	CLUPPIR	286-290
Hawaiian and Pacific Islander Owner	CLUPPIO	291-295
American Indian and Alaska Native Renter	CLUPIR	296-300
American Indian and Alaska Native Owner	CLUPIO	301-305
Asian Renter	CLUPAR	306-310
Asian Owner	CLUPAO	311-315
Hispanic Renter	CLUPHR	316-320
Hispanic Owner	CLUPHO	321-325
Black Renter	CLUPBR	326-330
Black Owner	CLUPBO	331-335
White and Other Renter	CLUPOR	336-340
White and Other Owner	CLUPOO	341-345
Renters	CLUPR	346-350
Owners	CLUPO	351-355
Blank		356-364
Demographic/Tenure group (code)	DTCODE	365-366
Demographic/Tenure group (label)	DTLABEL	367-368
Region	REGION	369-369
Division	DIV	370-370

Sample Size Input File: ACE2000_TB.FIN

Variable Description	Name	Places
Census Region	REGION	1
Census Division	DIV	2
State	STATE	3-4
Number of housing units budgeted for listing	BLIST	6-13
Target number of clusters for small sampling stratum	TCLUSTS	15-17
Target number of clusters for medium and large sampling strata	TCLUST	19-22
Target number of clusters in AIR stratum	TCLUSTA	24-26
Number of sampling strata in state	NSSINST	28
First ACE block cluster number for state	CSTART	30-34

Block Cluster Sampling Parameter File: ACE2000_PARABC.FIN

Variable Description	Name	Places
Census Region	REGION	1
Census Division	DIV	2
State code (01-72 = FIPS State Code)	STATE	3-4
Sampling Stratum	SS	5
Target number of block clusters	TCLUST	7-14
Total number of block clusters	NCLUST	16-23
Total number of housing units	NHU	25-32
First-step take-every	TE1	34-44
First-step random start	RS1	46-56