

# CONSTRUCTING THE CENSUS 2000 ADDRESS LIST

Robin A. Pennington, Miriam D. Rosenthal, U.S. Census Bureau  
Robin A. Pennington, U.S. Census Bureau, Washington, D.C. 20233

**KEY WORDS:** Master Address File; Housing Inventory; USPS Delivery Sequence File; Housing Unit Coverage; TIGER

**Abstract** One critical component of producing a complete count of persons and housing units in Census 2000 is a complete address list. In this paper, we review the construction of the Decennial Master Address File, the address list created for Census 2000. The following products and operations will be explained: Master Address File, Decennial Master Address File, United States Postal Service Delivery Sequence File, TIGER, geocoding, Address Listing, Block Canvassing, Local Update of Census Addresses, Mailout/Mailback, Update/Leave, List/Enumerate. The relative importance of each operation will be discussed in regards to the magnitude of the operation and its effect on coverage.

## Introduction:

Census 2000 is rapidly approaching. The decennial census is not just a count of people; it is also a count of housing units. People are associated with a geographic area. In order to count people in this fashion, the list of housing units must be accurate. For Census 2000, that housing inventory list is the Decennial Master Address File (DMAF). This is an extract of the addresses on the Master Address File (MAF) maintained by the U.S. Census Bureau. In this paper we discuss the creation of the MAF and the DMAF and their use in Census 2000.

Addresses that are included on the MAF come from a number of sources. Furthermore, the country is divided into areas according to the predominant address type, which determines the enumeration strategy; there are different operations that contribute to the address list in these different areas. The MAF contains a record for every housing unit that has been added to the list in any operation. The status codes from the operations determine whether a Census 2000 questionnaire will be delivered to the housing unit.

## Address Sources in Mailout/Mailback Areas:

Files and operations have been chosen for their potential to give a complete listing of housing units, to reflect changes in housing unit status and to yield information

about newly constructed housing units.

The majority of addresses in the country are in what is known for census purposes as the mailout/mailback area, which in general consists of areas with city-style addresses. A city-style address is of the type 121 Main Street. Most of the address list-building operations focus on these areas. The original source of addresses on the MAF for the mailout/mailback areas is the 1990 Census address file, the Address Control File (ACF). The first update to the ACF addresses is a United States Postal Service (USPS) Delivery Sequence File (DSF) of addresses. The USPS updates its DSF every month. The U.S. Census Bureau has procured some of these for the purpose of maintaining a current address list. The DSF lists both residential and non-residential addresses, with an indicator of which type each address is. Address information on the DSF comes from the local mail carriers, so the timeliness of address updates varies by carrier route.

The following paragraphs explain MAF construction, starting from the ACF and DSF address information. The ACF and DSF addresses are processed and assigned Census Bureau geographic codes. This processing occurs on a continuing basis over a number of years. Additional census operations augment and refine the MAF.

Until shortly before the census, the ACF addresses and the DSF residential addresses constitute the MAF. These addresses are tested against Census Bureau geographic information to determine their location at the census block level. The Census Bureau geographic information is maintained in the Topologically Integrated Geographic Encoding Referencing system (TIGER). This is a set of digitized maps for the entire country, as well as for some other regions in which the Bureau performs the census. Digitally connected to the maps for mailout/mailback areas are address ranges for street segments. This database was created around the time of the 1990 Census for the purpose of maintaining consistent and accurate census records.

Within the Bureau's Geography Division, there was an initial operation to assign address ranges to many of the streets in the TIGER files. An address range such as 100-198 is associated in the TIGER files with a street segment, by which is meant one side of a particular street

from one corner to the next. These often correlate with a ZIP+4 designation assigned by the USPS. When the MAF is processed in conjunction with the TIGER files, an address becomes coded to a specific block, where a block is usually represented by a polygon formed by the intersection of streets or streets and natural boundaries.

The coding of these addresses to the geographic level of a block is accomplished in two matching operations. The first is an automated matching operation. When an address on the MAF can be uniquely matched to the address range on a street segment that forms one of the boundaries of a particular block, the address is said to be geocoded to that block.

Addresses for which the appropriate range does not exist in the TIGER files, or for which there is more than one location given in TIGER, are not geocoded during the automated matching. These addresses then enter a second operation, clerical geocoding. Clerical geocoding takes place in all 12 Regional Census Centers. Trained geographic personnel use a variety of reference sources to find the correct location for the addresses and update the TIGER files so that these addresses geocode. Valid and geocoded addresses will appear on each address list used for a field operation.

For some of the addresses, there still is not enough information to locate them correctly in this clerical operation. The MAF indicates no associated block number for such an address. That address will not be included on any field operation address list because a block number must be known; if the housing unit with that address exists, it must be added to the list during the field operation.

There were two DSFs used to obtain addresses before the extant MAF was sent to a 100% Block Canvass field operation. In particular the November 1997 DSF and the September 1998 DSF were incorporated into the MAF. The Block Canvassing operation is the next major address list operation that was undertaken by the Census Bureau for Census 2000. It took place during the months of January - May, 1999 in the mailout/mailback areas of the country. In general housing units on the MAF that have been geocoded to a block are sent out for field verification in Block Canvassing. Added housing units are also anticipated. The Block Canvass listers receive materials for an assignment area, defined geographically. These materials are maps of the numbered blocks in the assignment area and the corresponding list of addresses coded to the specific blocks in the assignment area. Every address in an address register requires an action code from the Block Canvassing lister. The possible

action codes are ✓ for verify; C for a correction to the street name or directional, but not to the house number; D1 for delete; D2 for duplicate, implying the unit exists elsewhere on the list with a different, unmatchable designation such as a different street name or building name; U for uninhabitable; N for nonresidential. There is a Block Canvassing add page for the added housing units. There is also a Block Canvassing Special Place add page for domiciliary situations such as college dormitories and halfway houses. Special Places are enumerated in the census differently from housing units.

One action code that is not allowed due to the block-by-block canvass procedure of the Block Canvassing operation is a block number change for an address that is discovered to be in the wrong block. This can only be accomplished by deleting the unit from the incorrect block with an action code of D1 and adding the unit to the correct block. This may even be performed by two different listers, since the blocks can be in different assignment areas. In order to determine that the address was originally listed in the incorrect block, the addresses must be matched during the processing of the Block Canvassing data. The delete and add will then result in an action referred to as a geographic transfer of the unit.

Occurring in approximately the same time frame as Block Canvassing is a cooperative address list check with local governmental units (GUs) throughout the country. The Local Update of Census Addresses (LUCA) operation occurred in three phases. The phases are distinguished by the enumeration method to be used in the region. For the mailout/mailback areas, the operation is LUCA 1998. The LUCA 1999 operation for non-mailout/mailback areas will be discussed later in this paper. There is also a Supplemental LUCA operation conducted primarily in regions in which the enumeration technique designation changed after the LUCA 1998 areas were identified. This operation is a slight variant on the other LUCA operations and will not be discussed here. In LUCA 1998 the participating GUs received an address list and were asked for input mostly on added units but also on deleted units and corrected street names or directionals. The outcome of this operation is similar to that of Block Canvassing in that units are added to and deleted from blocks, and address corrections are made. Two distinctions between the operations are that the GUs do not necessarily derive the results through a field check but often by such procedures as referring to local address sources and current construction permits. Additionally, because the LUCA 1998 operation focused on changes to the address list, there was no verification of original addresses.

The vintage of DSF that is used for the files sent to LUCA 1998 and to Block Canvassing may differ. This leads to complications in determining the original source of an address as well as the status of an address. As one example of the complications, consider the implications of not incorporating the September 1998 DSF addresses into the list of addresses sent to a LUCA participant compared to incorporating the DSF addresses. The LUCA entity that receives its address list before the incorporation of the September 1998 DSF addresses will have to add any of the addresses that are new to this DSF. The Block Canvassing address list for this region will already have the DSF addresses included. Thus a LUCA 1998 add matches to a September 1998 DSF add, which is then presumably verified in Block Canvassing. The original source of this address on the file is not limited to one operation or file; both the DSF and LUCA 1998 picked this up as a new unit. If the LUCA entity receives an address list after the September 1998 DSF addresses have been incorporated, then LUCA 1998 will take no action on the address and Block Canvassing will verify the address, if it is correct. The DSF is the original source of the address in this case.

Consider the second case, in which both the LUCA entity and the Block Canvassing operation receive the same address list. If there is a new unit that has not yet appeared on the DSF, presumably both the LUCA 1998 operation, which depends on local knowledge, and the Block Canvassing operation, which takes place in the field, will add this unit to its address list. Both the LUCA 1998 and the Block Canvassing operations are credited with adding this address to the address list, and neither can be said to be the sole original source of this address.

A more complicated situation is when the status of a unit does not agree from one operation to another. In the case that LUCA 1998 and Block Canvassing are sent the same list of housing units, LUCA 1998 might add a unit that is not added in Block Canvassing. This will result in LUCA 1998 field verification of the unit. For the case when the addresses from LUCA 1998 are processed in time for inclusion in the Block Canvassing files, a unit added in LUCA 1998 that is deleted in Block Canvassing will result in a field verification of that unit.

If a unit changes residential status from one DSF to the next, the final unit status may be influenced by the timing of the operations. In the case that a unit is residential on the November 1997 DSF but has changed to nonresidential on the September 1998 DSF, a LUCA entity that receives an early address list will have this unit on its list; a LUCA entity receiving an address list incorporating the September 1998 DSF will not have this

unit on its list. It will not be on the address list sent to Block Canvassing either. When an address appears on a LUCA entity's address list, the LUCA participant can either delete the address, code it as nonresidential, make minor corrections or do nothing to it. When it does not appear on the LUCA entity's address list, the LUCA participant can only add the unit. If the unit is readded in either LUCA or Block Canvassing in this scenario, the status conflicts with the DSF status for that unit. The decision for inclusion or exclusion of the unit on the Census 2000 address list must take into account that there are different paths to the same outcome.

If we change the circumstances of the case just described slightly, so that the unit appears as residential on both DSFs, the set of possible final statuses changes, and the set of operation status paths that the unit may travel after such a DSF designation is vastly altered. The unit appears on the LUCA 1998 participant's list. If the LUCA participant denotes the address as a deleted or a nonresidential unit, there will be conflicting information on the address files. The LUCA delete action could match a delete action from Block Canvassing. In the case that the address list from LUCA 1998 is sent to Block Canvassing, the similar situation is that the unit is not readded in Block Canvassing. On the other hand, if a contradictory action is taken in Block Canvassing, then the LUCA 1998 status conflicts with both the DSF designation and the Block Canvassing action.

When there is conflicting information from different operations, there is a hierarchy of operations for the purpose of determining a unit's status. A Block Canvassing action has a higher priority than a LUCA 1998 action, and both Block Canvassing and LUCA 1998 have higher priority than the DSFs, so a unit that receives a delete action in LUCA 1998 but not in Block Canvassing, or is readded in Block Canvassing, will be included on the DMAF. Also from a housing unit coverage standpoint, it is risky to delete a unit from the list when only one operation has determined it to be a delete, particularly when the results of another operation conflict with that designation. In general when only one operation designates an address as non-existent or nonresidential, the address will still appear on the DMAF. If an operation subsequent to the creation of the DMAF designates a second delete, the unit will be flagged on the file as a delete.

#### **Address Sources in Update/Leave Areas:**

The initial DMAF was created in July and August of 1999 from criteria based on the action codes given in the operations that preceded its creation. This is the file of

addresses sent for questionnaire printing and labeling. The operations discussed thus far are the operations that occur before the creation of the initial DMAF in mailout/mailback areas. In addition to the operations described above, there are operations that add addresses in non-mailout/mailback areas.

After Mailout/Mailback, the second most common method of questionnaire delivery is Update/Leave. These are denoted as update/leave areas because the address list and maps are updated at the same time that questionnaires are delivered to each housing unit. There are fewer address list-building operations in update/leave areas than in the mailout/mailback areas. The DSF is not used to construct the address list in these regions because the addresses are primarily non-city-style. Instead, the address list for update/leave areas is constructed during a Census Bureau field operation called Address Listing. Census employees are sent to the field with maps of their assignment areas and are instructed to record the city-style address, non-city-style address or location description, or possibly some combination of the above, for every housing unit. In addition the location of the unit is noted on the census map. This operation took place in the fall of 1998.

At the completion of the processing of the Address Listing data, it is possible to tabulate the number of housing units in each block. Because the housing units in these areas may have nonstandard mailing addresses and may be recorded in census files solely with a location description, the GUs participating in the LUCA 1999 operation in these areas are sent lists of housing unit counts by block. When the LUCA 1999 participant disagrees with a Census block count, that block is sent out for LUCA 1999 recanvassing, in which census employees are redeployed to make updates to the address list. After processing the LUCA 1999 recanvassing materials, the block counts are retabulated.

In both the LUCA 1998 and LUCA 1999 operations, there is an appeal process for settling housing unit status or housing unit count discrepancies that are not resolved by the field verification process. Appeals will be resolved after the creation of the initial DMAF.

The operations described thus far yield the list of addresses sent to the contractors for the printing of the Census 2000 questionnaire address labels.

#### **DMAF Updates:**

There are a number of update operations that follow the creation of the initial DMAF. These updates to the

DMAF occur when addresses are added in operations up to Census Day, April 1, 2000. There will be November 1999, February 2000 and April 2000 DSFs adding addresses to the decennial census address files. This is an attempt to add newly constructed housing units to the list. The February 2000 DSF will contain the results of a concerted effort on the part of the USPS to update their files, called edit book week. The automated and clerical geocoding operations will take place on these address lists in Geography Division as before. These addresses will need to undergo special processing procedures in order for the housing units to receive questionnaires.

Another address update operation that occurs subsequent to the creation of the initial DMAF is the LUCA 1998 field verification and appeal process. As discussed above, many of the units receiving a conflicting status from the Block Canvassing and the LUCA 1998 operation will be sent for field verification by the Census Bureau; the results of the field verification will be sent to the GUs. At this stage it is possible for the GU to contest the Census Bureau's findings for particular units. At an appeal, the Census Bureau and the GU will submit their evidence of the status of a housing unit for independent review, and a ruling will be issued. Both the field verification and the appeal processes have the potential to change the status of a housing unit.

The last operation in mailout/mailback areas that adds addresses before Census Day is the New Construction operation, another cooperative effort with participating GUs. This operation uses the GU's local knowledge to identify new housing up until Census Day and takes place in February and March of 2000. Addresses added in this operation will also require special procedures for questionnaire delivery.

The last address list-building operation in the update/leave areas is the Update/Leave operation itself. This operation is responsible for having a census questionnaire hand delivered at every housing unit. In the process the MAF and the maps will be updated.

#### **Additional Enumeration Areas:**

In the most remote regions of the country, the housing units will be listed at the time of Census 2000. People will be enumerated concurrently. These operations are called List/Enumerate and Remote Alaska enumeration. This will be the only source of addresses in these regions.

Additionally there are special enumeration techniques for some regions of the country in which addresses were listed in previous operations. For example there is an

Urban Update/Leave operation for areas where mail delivery is considered to be problematic. The addresses have passed through all the operations of the mailout/mailback areas up until the time of the census, but the area will be visited by enumerators during the census, and, therefore, additions, deletions and corrections to the address list can be made.

**Conclusion:**

Ultimately the information from each operation will be fed back to the DMAF so that each housing unit record will contain a complete history of the actions taken in each operation. It will be possible to determine which operation(s) added the unit to the list. The primary operations responsible for adding addresses are the 1990 ACF, the November 1997 DSF, the September 1998 DSF, the November 1999 DSF, the February 2000 DSF, Block Canvassing, Address Listing, New Construction, Update/Leave, List/Enumerate and Remote Alaska, and LUCA 1998, LUCA 1999 and Supplemental LUCA. Because of the timing of events and the overlap of some of the operations, as well as the varying magnitude of the operations, it will not be possible to compare the operations for relative effectiveness in terms of numbers of addresses added to the MAF or deleted from the MAF or of numbers of corrections made to addresses. However with an understanding of the interrelatedness of the operations, some interpretation of the effectiveness of the operations may be attempted.

**Disclaimer:**

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau Publications. This report is released to inform interested parties of research and to encourage discussion.