



MEMORANDUM FOR      Distribution

From:                      Cynthia Clark  
                                 Associate Director for Methodology and Standards

Subject:                    An Evaluation of the Master Address File  
                                 Building Process

I am pleased to present the executive summary of one of the evaluation studies for the Census 2000 Dress Rehearsal. The dress rehearsal was conducted in three sites — Columbia, South Carolina; Menominee County, Wisconsin; and Sacramento, California. The evaluation studies cover detailed aspects of eight broad areas related to the census dress rehearsal — census questionnaire, address list, coverage measurement, coverage improvement, promotion activities, procedures addressing multiple options for census reporting, field operations, and technology.

The executive summary for each evaluation study is also available on the Census Bureau Internet site (<http://www.census.gov/census2000> and click on the link to “Evaluation”). Copies of the complete report may be obtained by contacting Carnelle Sligh at (301) 457-3525 or by e-mail at [carnelle.e.sligh@ccmail.census.gov](mailto:carnelle.e.sligh@ccmail.census.gov). Please note that the complete copy of the following reports will not be publically released: reports regarding procedures addressing multiple options for census reporting and the Evaluation of Housing Unit Coverage on the Master Address File.

The evaluations are distributed broadly to promote the open and thorough review of census processes and procedures. The primary purpose of the dress rehearsal is to simulate portions of the environment we anticipate for Census 2000, so we can identify and correct potential problems in the processes. Thus, the purpose of the evaluation studies is to provide analysis to support time critical review and possible refinements of Census 2000 operations and procedures.

The analysis and recommendations in the evaluation study reports are those of staff working on specific evaluations and, thus, do not represent the official position of the Census Bureau. They represent the results of an evaluation of a component of the census plan. They will be used to analyze and improve processes and procedures for Census 2000. The individual evaluation recommendations have not all yet been reviewed for incorporation in the official plan for Census 2000. These evaluation study reports will be used as input to the decision making process to refine the plans for Census 2000.

The Census Bureau will issue a report that synthesizes the recommendations from all the evaluation studies and provides the Census Bureau review of the dress rehearsal operation. This report will also indicate the Census Bureau’s official position on the utilization of these results in the Census 2000 operation. This report will be available July 30<sup>th</sup>.

# **An Evaluation of the Master Address File Building Process**

June 1999

---

Frank Vitrano/Lionel Howard  
Planning, Research, and  
Evaluation Division

## EXECUTIVE SUMMARY

The Master Address File building process for the Census 2000 Dress Rehearsal involved a series of operations that built on each other and ultimately resulted in the address list used to conduct the dress rehearsal. The Master Address File building process differed for areas with mail delivery to predominantly city-style addresses (mailout/mailback areas) and areas with predominantly non-city-style addresses (update/leave areas). During the implementation of the dress rehearsal, the plan for building the Master Address File for Census 2000 was re-engineered. Various operations were revised, deleted, or added. This evaluation reviews the steps in the Census 2000 Dress Rehearsal's Master Address File building process:

1. 1990 Address Control File
2. The Delivery Sequence File from the U.S. Postal Service
3. Targeted Multi-Unit Check
4. Targeted Canvassing
5. Local Update of Census Addresses
6. The Postal Validation Check
7. Urban Update/ Enumerate
8. Address Listing
9. Update/Leave
10. Be Counted/Telephone Questionnaire Assistance

This evaluation also looks at the final distribution of addresses on the Census Unedited File by the sources of the addresses.

The *1990 Address Control File* and the *United States Postal Service's Delivery Sequence File* were used to create the initial Master Address File for mailout/mailback areas of the country. The two files were matched against each other by ZIP Code and street name.

In the *Targeted Multi-Unit Check* operation, Geography Division compared the housing unit counts at multi-unit addresses (apartments, rooming houses, etc.) between the 1990 Address Control File and most recent Delivery Sequence File from the United States Postal Service. Where these counts differed, enumerators visited these basic street addresses to ensure that the census address list had the correct number of units. The enumerator also updated the unit designations. Note that this operation will not be conducted in Census 2000 because we will be doing a 100 percent block canvassing.

As a tool for targeted canvassing, the operation found fewer than 300 new housing units in both Sacramento and South Carolina, in the 19,377 and 12,304 housing units canvassed, respectively. In Sacramento, 31.2 percent of the multi-unit address count differences were resolved by phone. However, in South Carolina, only 15.5 percent were resolved by phone, requiring field visits for the vast majority of cases. *It is unclear from the Dress Rehearsal data whether improving coverage of housing units at multi-unit addresses can be done adequately over the phone.*

In the *Targeted Canvassing* operation, local officials identified and prioritized blocks where they expected hidden housing units to exist. Field staff looked for the missing or hidden units in the

particular block or subset of blocks identified by the local officials. Note that this operation will not be conducted in Census 2000 because we will be doing a 100 percent block canvassing.

As implemented, the Targeted Canvassing operation was productive in providing adds to the blocks canvassed in Sacramento and South Carolina. In Sacramento, the operation added 756 housing units during the canvassing of 19,477 housing units. In South Carolina, 111 housing units were added during the canvassing of 5,803 housing units.

During the *Local Update of Census Addresses* operation, local and tribal governments participated in a partnership program with the Census Bureau to conduct a review of the addresses on the Master Address File. Participating governments provided feedback in the form of recommended adds, deletes, or corrections of addresses to the Master Address File. The Census Bureau then conducted a series of operations to determine whether to accept or reject the recommended actions. The government again had the opportunity to review the Census Bureau's results and to provide additional feedback.

In terms of participation rates, the Census Bureau obtained the participation of the City of Sacramento and the Menominee Tribal government in the Local Update of Census Addresses program. In the South Carolina site, 31 of the 60 eligible governments (51.6 percent) participated. These government entities accounted for 98 percent of the 1990 Census housing units in the South Carolina site. *We recommend that the Bureau continue its efforts to form partnerships with local and tribal governments in this coverage improvement operation.*

The Local Update of Census Addresses operation varied in the capturing of new addresses, corrections to addresses, and addresses to delete across the sites. In Sacramento, 86.5 percent of the 4,528 corrections submitted, and 5.3 percent of the 2,918 adds submitted, were accepted in the operation (no deletes submitted). In Menominee, 100 percent of the 25 adds submitted, 97.6 percent of the 289 corrections submitted, and 60.7 percent of the deletes submitted, were accepted in the operation. In South Carolina, 43.2 percent of the 12,414 deletes submitted, 56.3 percent of the 26,983 corrections submitted, and 12.6 percent of the adds submitted, were accepted by the operation. It should also be noted that there was a large rejection rate of initial submissions across the sites. *We recommend that the Census Bureau do more to improve the process, and to educate and train Local Update of Census Addresses participants to make this coverage improvement operation more efficient.* Based on Dress Rehearsal experience revisions to the Local Update of Census Addresses program have been made to improve both the process and training for Census 2000.

In the *Postal Validation Check* operation, United States Postal Service employees verified the completeness of the Master Address File by comparing Master Address File addresses with the addresses in their carrier delivery routes. The Census Bureau limited the operation to 29 ZIP Codes (seven in South Carolina and 22 in Sacramento) that were entirely within the dress rehearsal sites and entirely inside mailout/mailback areas. The operation's primary purpose was to capture late new construction in time for the mailout of census questionnaires.

The Postal Validation Check operation provided a substantial number of addresses recommended for deletes. In Sacramento, 75.7 percent of the 12,551 addresses paid for were deletes. In South Carolina, 67.3 percent of the 4,856 addresses paid for were deletes. With existing addresses

already printed on census questionnaires and ready for delivery, we could not make use of the address deletions or corrections that are a standard part of the Postal Service's procedure. The United States Postal Service and Census Bureau define valid housing units differently, so we would not process the deletes even if they were known earlier. Because of the timing of the operation and the inconsistency of how housing units are defined, we are only able to make use of new addresses in the operation. However, we pay the Postal Service for the deletes. The Postal Service is resistant to changing their standard procedure. The Census Bureau recognizes that requested exceptions to the procedure could result in errors.

There was a high match rate between new addresses provided by the Postal Service and addresses we already had on the Master Address File. In Sacramento, 41.2 percent of the 3,189 adds matched to the Master Address File. In South Carolina, 53.8 percent of the 1,223 adds matched to the file. Therefore, having a system to identify duplication before questionnaire delivery was an important component of the dress rehearsal process for the Postal Validation Check. *We recommend that as in the dress rehearsal, systems be put into place to look out for duplicate addresses provided in the Postal Validation Check operation during Census 2000. Given the expense of processing cards we cannot use and the high match rate of add cards to addresses we already have, a cost benefit analysis of the Postal Validation Check operation should be conducted as part of Census 2000.*

***Urban Update/Enumerate*** was conducted in six municipalities in the South Carolina site that did not participate in the Local Update of Census Addresses operation and had a high concentration of post office box addresses. Field staff visited these areas to update the address list and to collect census data for all newly discovered units at the time of visit. Unfortunately, the adds that came from the operation were processed along with nonresponse follow-up adds. Therefore, the impact of the operation cannot be examined. Note that this operation will not be conducted in Census 2000 because we will be doing a 100 percent block canvassing operation. *We recommend that additional flags be added to the Master Address File to more clearly show the history of each address through the various building operations.*

In ***Address Listing***, census enumerators went door-to-door to identify the mailing address and physical location of housing units in areas where the Census Bureau did not have good locatable addresses from either the 1990 Address Control File or the Delivery Sequence Files provided by the United States Postal Service. The enumerators also map-spotted each housing unit on a block map. In Menominee, listers were able to obtain a mailing address at 51.6 percent of the 2,060 addresses listed. Of the 1,063 housing units where they obtained a mailing address, 91.5 percent were non-city style addresses. In South Carolina, listers were able to obtain a mailing address at 75.9 percent of the 66,704 addresses listed. Of the 50,595 where they obtained a mailing address, 39.1 percent non-city style address.

The ***Update/Leave*** operation was conducted just prior to Census day. Enumerators canvassed each block in their assignment area, matching what was found on the ground to the list of addresses in their register. They updated the register by adding new addresses, deleting addresses they could not locate or that were duplicates, and correcting addresses. The added units were also map spotted, and the appropriate form type was addressed.

The Update/Leave operation provided 4,331 new addresses, 7,453 corrections to existing

addresses, and 4,225 deletes in South Carolina; 96 new addresses, 566 corrections to existing address, and 87 deletes in Menominee.

The *Be Counted and Telephone Questionnaire Assistance* programs are two ways that people can complete a census form if they were not otherwise enumerated. The Be Counted program gave residents access to questionnaires in their local community, while the Telephone Questionnaire Assistance program collected census data over the phone. In both programs, if the address could be geocoded and it was not already in the census address inventory it was added to the Master Address file.

In Sacramento, 84.3 percent of the 1,575 Be Counted cases were successfully geocoded. However, only 68.3 percent of the 1,328 geocoded cases were geocoded in time for inclusion in the Census. In South Carolina and Menominee, 91.7 and 76.9 percent of the 661 and 13 geocoded cases, respectively, were geocoded in time to be included in the Census. *More work needs to go into planning the geocoding of Be Counted Addresses in order to make sure that all can be successfully geocoded and matched to the Master Address File in time for inclusion in the census.*

The relative impact of each operation on the building of the Master Address File could not be adequately assessed in the evaluation. This was largely due to the manner in which data were retained on the Master Address File extracts used in the dress rehearsal. In particular, we could not obtain the universe of addresses going into each operation. The universe of addresses going into each operation would have provided a base against which to measure the relative impact of the operation. Additionally, the Master Address File extracts only retained the results of the most recent field operation. For example, if an address came in through the Targeted Canvassing operation, and later through the Postal Validation Check operation, the Master Address File would flag the Postal Validation Check operation as the address' input source. By updating the file with the most recent field operation, it was not possible to determine which operation was the initial input source. With these limitations, we still attempted to gain some sense of each operation's relative impact by examining the final status of housing units on the Census Unedited File. The Census Unedited File was created from the Master Address File extracts.

We found that more than 100 combinations of Master Address File sources appeared on the Census Unedited File when examining the *final distribution of addresses on the Master Address File by source*. Of the 149,477 valid addresses, in Sacramento, 89.7 percent appeared on both the Delivery Sequence File and the 1990 Address Control File. Of the 7,955 deleted addresses in Sacramento, 25.1 percent appeared on both files. The Delivery Sequence File and the 1990 Address Control File also provided 58.2 percent of the 210,164 valid addresses in the mailout/mailback portions of South Carolina. And, 14.9 percent of the 17,817 deleted addresses in the mailout/mailback portions of South Carolina. By representing such a large percentage of the good addresses and such a small percentage of the deletes, the combination of the Delivery Sequence File and the Address Control File were good initial sources for the building of the Master Address File in mailout/mailback areas. Beyond the Address Control File and the Delivery Sequence File, it was difficult to examine the impact of each Master Address File source.

*For Census 2000, we recommend a thorough review of the flags set on the Master Address File that show the relative impact of each operation. In particular, the creation of specific variable*

*fields that follows the history of an address, as it relates to an operation. History refers to the impact (type of action code suggested - add, delete, corrected, move to a different block, etc.) of various operations on an address throughout the building process. Fields should also be created to determine if an address was part, or not part, of the initial universe for specific operations.*

The Decennial Statistical Studies Division is currently working with Geography Division and the Decennial Systems and Contracts Management Office to make these improvements. The availability of cost information, along with a revised set of flags, will allow us to do a thorough cost benefit analysis of the entire Master Address File Building Process.