# Analysis of Deleted and Added Housing Units in Census 2000 Measured by the Accuracy and Coverage Evaluation

FINAL REPORT

This evaluation reports the results of research and analysis undertaken by the U.S. Census Bureau.  It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning.  Findings from the Census 2000 TXE Program reports are integrated into topic reports that provide context and background for broader interpretation of results.

Damon Smith
Diane F. Barrett
Michael Beaghen
Decennial Statistical
 Studies Division

U S C E N S U S B U R E A U
*Helping You Make Informed Decisions*

# CONTENTS

# List Of Tables

# EXECUTIVE SUMMARY

This evaluation examines the housing units added and deleted in Census 2000 using the results of the 2000 Accuracy and Coverage Evaluation survey. It evaluates the changes in the Census 2000 housing unit inventory between the creation of the January 2000 Decennial Master Address File and the Hundred Percent Census Unedited File Prime. The study analyzes the Decennial Master Address File deletes and determines which were correctly deleted or deleted in error by the census, giving particular attention to two census operations that deleted addresses, the Housing Unit Unduplication Study and the Kill Process. Likewise, the study analyzes the Hundred Percent Census Unedited File adds and determines which were correctly added or added in error, according to the Accuracy and Coverage Evaluation. The Initial Housing Unit match and Person Interview of the Accuracy and Coverage Evaluation afforded an opportunity to examine the enumeration status of housing units that were later deleted from the census inventory. The Final Housing Unit match of the Accuracy and Coverage Evaluation afforded an opportunity to examine the enumeration status of housing units that were added to the census inventory and gave some insight to the census operations that added addresses.

**How were the deleted census housing units identified?**

We identified census deleted addresses as those that were on the January 2000 Decennial Master Address File and not on the Hundred Percent Census Unedited File Prime[1] . These addresses are called Decennial Master Address File deletes in this report. We classified the Decennial Master Address File deletes as either "correctly deleted" or "deleted in error" from the census address inventory.

We classified a deleted census housing unit as "correctly deleted" if the Accuracy and Coverage Evaluation survey identified the unit as an erroneous enumeration or as a duplicate. That is, a delete was counted as correctly deleted if the housing unit was coded in the Accuracy and Coverage Evaluation as not a housing unit, geocoding error, unresolved housing unit status, or a duplicate. Furthermore, a delete was counted as correct if an Accuracy and Coverage Evaluation housing unit that was matched to it was either deleted based on the results of the survey's Person Interview or matched to a different census unit in the survey's Final Housing Unit match. If the Accuracy and Coverage Evaluation did not determine a unit to be a duplicate but the census did in its Kill Process or surviving Master Address File ID process then we presumed the Accuracy and Coverage Evaluation did not find the duplicate because it was outside the Accuracy and Coverage Evaluation search area and counted the delete as correctly deleted. We did

---

[1] The Hundred Percent Census Unedited File (HCUF) Prime was the second of three versions created. The file contained flagged potential duplicate housing units, which were later deleted if further research deemed the unit to be a duplicate.

not assume that units deleted in the Housing Unit Unduplication Operation were correctly deleted.

In some situations census units may have been deleted because the associated household of people was identified as duplicated, but this does not imply that the housing unit was duplicated (Overview of the Duplicate Housing Unit Operations, Nash 2000).  At the same time we point out that the estimates of incorrect deletions by the Housing Unit Unduplication Operation represent an upper bound because the Accuracy and Coverage Evaluation could not identify duplicates outside its search area.  Some of these deletes could have had a duplicated housing unit outside the search area.

Addresses "deleted in error" were defined as those units where the Accuracy and Coverage Evaluation confirmed the housing units to exist on Census Day and found no evidence of duplication.  Any of the following scenarios would have provided evidence of duplication: a duplicate was identified by the Accuracy and Coverage Evaluation, or the deleted unit was identified as a duplicate by census field operations, or the Accuracy and Coverage Evaluation housing unit was matched to a census unit that was a duplicate of another census record identified outside the Accuracy and Coverage Evaluation search area.

**How were the added census housing units identified?**

Using the same files (January 2000 Decennial Master Address File and Hundred Percent Census Unedited File Prime), we identified census adds as addresses that were on the Hundred Percent Census Unedited File Prime and not on the January 2000 Decennial Master Address File.  These addresses are called Hundred Percent Census Unedited File adds in this paper.  We classified the Hundred Percent Census Unedited File adds as either "correctly added" or "added in error" to the census address inventory.  "Correctly added" addresses were defined as those units that were confirmed to exist as housing units on Census Day by the Accuracy and Coverage Evaluation.  Addresses "added in error" were defined as those units that were added to the Census 2000 address inventory but were determined by the Accuracy and Coverage Evaluation not to exist as housing units on Census Day, to be duplicates, or to be geocoding error.

**How were the Accuracy and Coverage Evaluation results used to assess the deleted and added units?**

Results from the Accuracy and Coverage Evaluation survey were used to assess the validity of units added and deleted by the census.  The assumption we made was that the Accuracy and Coverage Evaluation survey results were reliable because there were several opportunities during the survey operations to verify the status of the units.

A limitation of the study was that the selection of January 2000 as a reference point to define adds and deletes was somewhat arbitrary. January 2000 was selected as the reference point because that was the latest census inventory of addresses available at the time of the Initial Housing Unit match that took place between the survey and the census. Adds and deletes were continuously processed for the census after January 2000. If instead we had been able to match the survey results to a later version of the Decennial Master Address File (for example the February 2000 or March 2000 Decennial Master Address File), we would likely have had fewer units identified as adds and deletes in this study. But January 2000 was selected out of consideration for the Accuracy and Coverage Evaluation survey operations, not with consideration as to what date would be best to use as a reference point for defining census adds and deletes.

**How does the Census 2000 Housing Unit Unduplication Operation relate to this study?**

It is also very important to note how our study relates to the Census 2000 Housing Unit Unduplication Operation with respect to the **reinstated** units by census operations and in this analysis. Before the start of the 2000 Accuracy and Coverage Evaluation person matching survey operations, the census flagged addresses it thought to be potential duplicates and removed them from the existing address inventory. The resulting census address inventory file was the file that was used in all future survey processing. However, some of the initially deleted units were later reclassified as actually existing as housing units on Census Day. Those units were *reinstated* and added back to the final census address inventory. But, because of timing, the Accuracy and Coverage Evaluation could not go back to include the reinstated units in its processing. As a result, reinstated units are not included in this analysis. Although the census files included these reinstated units, we removed them from any data files used for this analysis. However, the deletes from the Housing Unit Unduplication Operation were evaluated in this study.

**Conclusions**

Being mindful of these limitations, regarding the exclusion of reinstated units from this analysis, the somewhat arbitrary nature of the January 2000 reference point in identifying census adds and deletes, and the inability of the Accuracy and Coverage Evaluation to detect duplicates outside its search area, the major conclusions[2] of this study are as follows:

---

2  The results presented in this paper are representative of housing unit counts rather than population counts.

**Most (85.6 percent) of the housing units census deleted were correctly deleted.** That is, the Accuracy and Coverage Evaluation was in agreement with the census that the address did not exist as a housing unit or was a duplicate on Census Day. Of the correctly deleted census housing units, 63.1 percent were not a housing unit or did not exist on Census Day according to the Accuracy and Coverage Evaluation. Accuracy and Coverage Evaluation duplicate addresses accounted for a total of 12.8 percent of the correctly deleted Decennial Master Address File deletes.

**About 14.4 percent of the Decennial Master Address File deleted units were deleted in error according to the Accuracy and Coverage Evaluation.** This represents a weighted estimate of 1.2 million housing units which may have been deleted in error by the census. The Accuracy and Coverage Evaluation confirmed the housing units to exist on Census Day and determined there was no evidence of duplication. These units were visited by the Accuracy and Coverage Evaluation interviewers during the Initial Housing Unit Followup operation (which took place between February 12, 2000 and April 4, 2000) or the Person Interview (which took place from June 19, 2000 through August 18, 2000) to determine if the addresses existed as a housing units on Census Day.

These Decennial Master Address File deletes that the census deleted in error contributed substantially to census omissions. The Census 2000 Housing Unit Coverage Study, Evaluation O.3, found the national estimate of omissions was 3.62 percent or about four million P-sample nonmatches (Barrett et al, 2003).

**Of the census operations that excluded units from the census, the Kill Process (61.3 percent) and the Housing Unit Unduplication Operation (29.3 percent) accounted for 90.7 percent of the 1.2 million erroneous deletes. Even so, the Kill Process deleted units correctly most of the time (89.7 percent) and the Housing Unit Unduplication Operation correctly deleted units more than half of the time (64.0 percent).** Note that the estimate for the erroneously deleted units by the Housing Unit Unduplication Operation is an upper bound because of the limitations of the Accuracy and Coverage Evaluation search area. Some of these units could have been duplicated housing units outside the search area. If the census deleted the correctly geocoded unit, which was inside the search area, while keeping the incorrectly geocoded unit, then our study classified the unit as erroneously deleted.

**Most (83.9 percent) of the housing units Census 2000 added were found to be correct enumerations by the Accuracy and Coverage Evaluation.** Results from the Accuracy and Coverage Evaluation Final Housing Unit phase found that these units existed as housing units on Census Day and were neither duplicates nor geocoding error.

**About 16.1 percent of the Hundred Percent Census Unedited File adds were added in error by the census.** The Final Housing Unit Phase of the Accuracy and Coverage Evaluation classified these units as erroneous enumerations; that is, they either were not a housing unit on Census Day, were duplicates of another address, or were geocoded to the incorrect Accuracy and Coverage Evaluation cluster.

**Recommendations**

The insights this evaluation yielded and the questions it raised have led to the following recommendations:

- Conduct research on the census Kill Process so that it deletes fewer valid housing units. Its error rate was small as a percentage, but there were enough incorrect deletions to make efforts at reduction important. This evaluation found that the Kill Process deleted many of the housing units that should have been included in the census because addresses were undeliverable as addressed by the United States Postal Service and were deleted from the address inventory during Nonresponse Followup. For an address to be deleted, at least two census operations had to confirm the address as a delete. We recommend further research be conducted that evaluates the use of the "undeliverable as addressed" addresses as one of the determining factors for a kill.

- Conduct a housing unit coverage study for the 2010 Census but design it with more emphasis on evaluating census processes.

- Obtain better estimates of gross census error as opposed to concentrating on net coverage error. For example, for the Accuracy and Coverage Evaluation housing unit followup, have a consistently applied and wide search area. This will allow us to better identify census geocoding error and distinguish geocoding error from the housing units that do not exist, and also to identify housing unit duplication that is beyond the Accuracy and Coverage Evaluation search area as defined for 2000.

- Conduct research to refine procedures for identifying and removing units that are believed to be duplicate housing units. This evaluation found that the Housing Unit Unduplication Operation may have deleted many units which should have been included in the census. Planning and work has already begun on building a housing unit unduplication process in the 2010 Census.

- Conduct further research that examines the breakdown of housing unit status by census field operations, with emphasis on status obtained in Nonresponse Followup and Coverage Improvement Followup Operations.

# 1. BACKGROUND

The Accuracy and Coverage Evaluation (A.C.E.) measured the overall and differential coverage of the United States population and the housing unit inventory in Census 2000. Housing units within the A.C.E. sample clusters were listed independently of the census in the Fall of 1999. This listing is called the independent listing. It was matched to the January 2000 version of the Decennial Master Address File (DMAF). The DMAF was a system of files and linkages that contained the current address inventory, enumeration and processing status of housing units which were necessary to support many Census 2000 operations . The A.C.E. and census housing units were matched in order to identify the housing units not found on both listing inventories. If the units did not match, they were called a nonmatch (these units include possibly matched units). The nonmatches were reconciled with a field interview. The field interviews were called the housing unit followup (HUFU). HUFU interviewing was conducted to gather more information about the housing units in question to resolve differences between the A.C.E. and census listing. The match and reconciliation were called the Initial Housing Unit (IHU) match. From the IHU match, a list of A.C.E. housing units confirmed to exist within the block clusters was prepared . Subsequently, there was the large block cluster subsampling of the independent listing units giving us the Population sample or P sample, and the subsample of census listing of housing units in A.C.E. areas, yielding the Enumeration sample or E sample. The P sample consists of addresses enumerated independent of the census. The E sample consists of addresses enumerated in the census.

Matching and reconciliation took place for housing units in two stages, the IHU phase just described and the Final Housing Unit (FHU) matching phase. The FHU operation processed the updates in the census address inventory that occurred between the creation of the January 2000 DMAF and the creation of the Hundred Percent Census Unedited File Prime (HCUF Prime)[3]. The FHU operation also processed the updates due to the large block subsampling, which occurred in clusters with a large number of A.C.E. units; the E-sample identification, where A.C.E. units that were not in the P sample were removed from further processing and the census units that were not in the E sample are assigned an E-sample indicator of 2 (those units in the E sample were assigned an E-sample indicator of 1); and the A.C.E. person interviewing, where FHU computer processing resolved addresses based on the person interview.

Then, in the FHU clerical matching phase we processed housing units that were added and deleted to the census list after the Initial Housing Unit matching operation. A.C.E.

---

[3]  The HCUF contains enumeration data and processing status information which are necessary to support the Census 2000 operations. There were three versions of the HCUF created, which were noted in the HCUF documentation (Phillip 2002). The HCUF Prime was the second version created and the version A.C.E. used.

addresses confirmed not to exist on Census Day during A.C.E. Person Interviewing (PI) were deleted from the A.C.E. inventory of addresses. These addresses were called Computer Assisted Person Interview removed or CAPI removed. Updated address information was also collected during the person interview. Only the updates in the HCUF and the A.C.E. inventory required clerical review in the FHU match. Matching and field interviews were also a part of the FHU match phase, just as in the initial housing unit match phase discussed earlier.

The census records used for IHU matching to the P sample were derived from the census inventory of addresses available on the January 2000 DMAF. The FHU match, however, used the census inventory of addresses that existed on the HCUF Prime for matching.

There were some records on the January 2000 DMAF extract that were not on the HCUF Prime Census inventory of housing units, which we refer to here as DMAF deletes[4] . These DMAF deletes were removed from the Census 2000 address inventory and were not processed in the Final Housing Unit phase of the A.C.E. operation. As a result of deleted census units (DMAF deletes), A.C.E. housing units that matched to those deleted census units, in IHU match, became new nonmatches in the FHU match phase.

There were also some records on the HCUF Prime housing unit inventory that were not on the January 2000 DMAF extract, which we refer to as HCUF adds[5]. The FHU computer processing identified these HCUF adds for clerical review. HCUF adds were processed in the FHU operation and were eligible to be matched to A.C.E. housing units.

This paper evaluates the changes in the Census 2000 housing unit inventory between the creation of the January 2000 DMAF and the HCUF Prime. The January 2000 DMAF was used in this analysis because this was the version of the DMAF that A.C.E. used for the initial housing unit match. This choice of date restricts the universe of cases that we examine in our study. Units deleted from the census inventory before January 1, 2000 are not included in this study, as are units added after January 1, 2000 and subsequently deleted. Similarly, units added before January 1, 2000 are not included in the evaluation of adds.

The study analyzes the DMAF deletes and determines which were correctly deleted or deleted in error by the census. Likewise, the study analyzes the HCUF adds and determines which were correctly added or added in error. The IHU match and Person

---

[4]    The use of the words "DMAF deletes" are particular to this paper and may not correspond exactly to the use of the word in other Census 2000 contexts.

[5]    The use of the words "HCUF adds" are particular to this paper and may not correspond exactly to the use of the word in other Census 2000 contexts.

Interview afford an opportunity to examine the enumeration status of housing units that were later deleted from the census inventory.  The FHU match affords an opportunity to examine the enumeration status of housing units that were added to the census inventory and gives some insight to the census operations that added addresses.  An important related paper is the Housing Unit Coverage Study (HUCS), which evaluates Census 2000 housing unit coverage (Barrett et al, 2003).  Another important related paper is 'An Assessment of Addresses on the Master Address File "Missing" in the Census or Geocoded to the Wrong Collection Block' (Ruhnke, 2003).

# 2. METHODS

## 2.1 DMAF Deletes

The DMAF deletes were identified by comparing the A.C.E. extract of the January 2000 DMAF to the Hundred Percent Census Unedited File (HCUF) Prime files. Units that were on the January 2000 DMAF and not on the HCUF Prime address inventory were called DMAF deletes.

*Exclusion of the Census Reinstated Units from the Analysis* - Before the start of the 2000 A.C.E. person matching, the census flagged addresses it thought to be potential duplicates and removed them from the existing address inventory. Later, after the adds and deletes were identified and processed from the various census coverage improvement operations, some of these deleted potential duplicates were **reinstated** and added back to the final address inventory . Although the census files included these reinstated units, we removed them from any data files used for this analysis. Because of timing, the A.C.E. was not able to include the reinstated units in its processing. Therefore, they are not among the "adds" or "deletes" we analyze in this paper. However, the deletes from the Housing Unit Unduplication Operation are evaluated in this study.

## 2.2 Correctly Deleted Housing Units

We classified a deleted census housing unit as correctly deleted if the A.C.E. identified the unit as an erroneous enumeration or as a duplicate. Thus a delete was counted as correctly deleted if the housing unit was coded in A.C.E. as not a housing unit, geocoding error, or a duplicate. Furthermore, a delete was counted as correct if an A.C.E. housing unit that was matched to it was either removed by the A.C.E. based on the results of person interview or if the remaining A.C.E. nonmatched unit matched another census unit in the A.C.E. Final Housing Unit match. If the A.C.E. did not determine a unit to be a duplicate but the census did in its Kill Process or surviving MAFID process then we presumed the A.C.E. did not find the duplicate because it was outside the A.C.E. search area and counted the delete as correctly deleted. We did not assume that units deleted in the Housing Unit Unduplication Operation were correctly deleted, because census units may have been deleted because of a household of people being identified as duplicated, not necessarily that the housing unit was duplicated. See the section of limits for more discussion on the evaluation of deletes from the Housing Unit Unduplication Operation.

*2.2.1 DMAF deletes identified as correctly deleted by A.C.E. Initial Housing Unit Match*

The DMAF deletes determined to be erroneously enumerated in the Initial Housing Unit (IHU) match and followup of the A.C.E. were classified as correctly deleted in this study. We also counted those with unresolved enumeration status from the IHU as correctly deleted. Below are the codes from the A.C.E. IHU match operation that were counted as correct deletions.

- *Geocoding Error or GE* - The census housing unit existed as a housing unit at the time of the A.C.E. followup interview, but was incorrectly listed in the A.C.E. block cluster. This housing unit was erroneously enumerated in the A.C.E. block cluster because of a geocoding error.

- *Census Duplicates or DE* - In A.C.E. there were two identifiers assigned to duplicate records within the E sample. The unit could be a primary or duplicate. When A.C.E. encountered one or more duplicates of a particular housing unit, one of them was chosen to be the primary unit and coded as a non-duplicate, while the others were coded as duplicates. In this study a DMAF delete coded as either a primary to a duplicate or as a duplicate were classified as correct deletes. These two groups were:

  - ▸ DMAF deletes that were coded DE.
  - ▸ DMAF deletes coded by A.C.E. as a correct enumeration and linked to one or more duplicate units (primary to duplicates).

- *Not a Housing Unit or EE* - According to A.C.E., the census address was nonresidential or did not exist on Census Day; that is, the address was for a group quarters, business, or the unit was demolished, burned down, uninhabitable or could not be located.

- *Unresolved Units or UE, MU* - These were cases where there was not enough information on the A.C.E. followup form to assign a code to the census nonmatched housing unit. The assigned unresolved A.C.E. match codes were UE and MU. UE match codes were assigned when there was not enough information on the A.C.E. housing unit followup form to assign a code to the census nonmatched housing unit with certainty. MU match codes were assigned when the A.C.E. and census addresses matched and there was not enough information on the housing unit followup form to confirm the match as a housing unit with certainty. The followup interview for these cases (UE, MU) was not done, was incomplete, was never sent, had contradictory information, or was a noninterview.

### 2.2.2  DMAF deletes identified as correctly deleted by A.C.E. person interviewing

A DMAF delete that was matched in IHU to an A.C.E. unit was identified as a correct deletion if the A.C.E. person interview determined the matching A.C.E. address was not a housing unit on Census Day. The reasoning is that the matched A.C.E. unit represented the same address as the census unit. After the IHU match operation, a list of addresses confirmed to be housing units in the A.C.E. sample block clusters on Census Day was created. Clusters with a large number of housing units were subsampled creating the enhanced list. The enhanced list was used for A.C.E. person interviewing. If the results from the person interview, that is, the computer assisted person interview or CAPI, indicated the housing unit did not exist as a housing unit on Census Day, then the matched census unit was counted as correctly deleted.

### 2.2.3  DMAF deletes identified as duplicates by the Kill Process

A unit coded as a correct enumeration by the A.C.E. was considered correctly deleted if the census identified the unit as a duplicate in the Kill Process. The Kill Process was a census operation that identified MAFIDs that most likely did not uniquely identify housing units as of Census Day (Treat, 2000). A total of approximately 9.0 million housing units were deleted during the Kill Process. In this study, we classified as units correctly deleted those deletes where A.C.E. confirmed the unit to exist on Census Day but the unit was deleted by the Kill Process because it was identified as a duplicate.

### 2.2.4  Duplicate records identified on the Master Address File or Surviving MAFIDs

During census processes, surviving MAFIDs were assigned to records on the MAF where duplicate units were identified.  In this study, we classified DMAF deletes where A.C.E. confirmed the unit to exist on Census Day but the unit was identified by the census as a surviving MAFID or duplicate of another record as correctly deleted DMAF deletes. The surviving MAFID indicated that there was another record on the MAF that represented the same unit, that is, a duplicate.

### 2.2.5  Deletes identified as correctly deleted by the Final Housing Unit Match

A DMAF delete that had been matched in IHU could be identified as a correct delete based on what happened to the matching A.C.E. unit in the FHU match. If the matching A.C.E. unit was matched to a census unit in the FHU match operation the DMAF delete was classified as a correct deletion. An A.C.E. unit that had matched in IHU to a deleted census unit could later match to a different census housing unit listing, that is, one with a unique census ID, for several reasons. First, the deleted census unit was a duplicate in IHU and the A.C.E. matched the non-deleted duplicate. Second, the same housing unit, with a unique census ID, could have been added back in the census in an operation such

6

as Update/Leave (U/L). In either case we classify the DMAF delete as correctly deleted. It is necessary to consider if the A.C.E. unit matched in FHU because a duplicate census unit may not have been identified as a duplicate by A.C.E. operations. This could happen if the duplicate had an address that was in some respect poor or ambiguous.

Also, in a small number of cases if it was determined in FHU followup that the matching A.C.E. unit was not a housing unit on Census Day then the DMAF delete was counted as correctly deleted.

An A.C.E. housing unit matching in IHU to a DMAF delete is identified as matched in FHU if its FHU match code is M. It is identified as removed from the A.C.E. if its FHU match code is ZI, GI or DI .

- *Match or M* - The P-sample and E-sample housing units matched in A.C.E.'s Final Housing Unit match operation.

- *Removed or ZI* - The P-sample address was incorrectly included in the A.C.E. list of housing units. The code was also used when the P-sample address did not refer to a housing unit at the time of the followup interview. For example, the housing unit burned or the mobile home moved. Another example, the address was commercial property or a special place.

- *Geocoding Error or GI* - The P-sample housing unit existed as a housing unit at the time of the followup interview, but was incorrectly listed in the block cluster. The housing unit was an A.C.E. geocoding error.

- *Duplicated or DI* - The housing unit should not have been listed in the P sample. This address was a duplicate of another P-sample address. This address was removed from the P sample.

Results from A.C.E. and census were used to derive the estimates for the correctly deleted DMAF deletes. Table 1e shows how we classified the correctly deleted DMAF deletes by the various A.C.E. and census delete reasons. Many of the reasons overlapped with one another, for example, a unit could be identified as a surviving MAFID as well as a CAPI removed, or a unit was identified by A.C.E. as not a housing unit (EE) as well as a kill, and so on. So we developed an order or hierarchy of factors to assign units with overlapping reasons. Therefore, estimates for the correctly deleted DMAF deletes were derived using the hierarchy listed below. As a result, the totals assigned for the categories represent the subset of that category that does not overlap with the higher priority categories.

- A.C.E. housing units erroneously enumerated in the IHU match (GE, DE, EE, UE, MU)
- Census duplicates (surviving MAFIDs)
- Housing units deleted as a result of the A.C.E. person interview (CAPI removed)
- A.C.E. primary to duplicates
- Census duplicates in the Kill Process
- A.C.E. housing units that matched to another census unit in the FHU match
- A.C.E. housing units removed from the P sample as a result of the FHU match

## 2.3    Deleted in Error Housing Units

We classify DMAF deletes as deleted in error if they were confirmed to exist on Census Day by A.C.E. and they were not determined to be duplicates as described in the previous section.  A census listing could be confirmed to exist as a housing unit on Census Day in either of two ways:  by the IHU field followup if it was a nonmatch in IHU match, or by the A.C.E. person interview (CAPI), if it was matched to an A.C.E. unit in IHU.  A more detailed description of these two ways follows.

- *Nonmatch or CE* code in IHU: the census housing unit existed as a housing unit at the time of the A.C.E. housing unit followup interview in March 2000 and was correctly geocoded in the block cluster.  The census address was not matched to an A.C.E. address .  These housing units were counted as incorrectly deleted unless they were in the following groups described in greater detail in the previous section:

  - Housing units that were identified as surviving MAFIDs,
  - Units identified as an A.C.E. primary to duplicates,
  - Duplicates identified from the Kill Process.

- *Match or M* code in IHU, and the matching A.C.E. unit has a CI code in FHU: the census housing unit was matched to an A.C.E. housing unit in the IHU match phase of A.C.E.  The A.C.E. housing unit was confirmed to exist as a housing unit on Census Day by the A.C.E. person interview (CAPI) and was not matched in FHU; its FHU code is CI.

  Thus the DMAF deletes that matched A.C.E. units in IHU were counted as incorrectly deleted unless they were in the following groups:

8

- Units identified as surviving MAFIDs,
- Units identified by the A.C.E. as primaries to duplicates,
- Duplicates identified in the Kill Process,
- Units that were evaluated in the FHU match operation that were determined to not be housing units on Census Day,
- A.C.E. units deleted as a result of the A.C.E. person interview (CAPI removed),
- Units matched in the FHU match or removed from A.C.E. in FHU.

## 2.4  Census Operations that deleted addresses from the inventory

There were four operations that deleted units from the census. The two major processes that identified the majority of the housing units on the DMAF that were deleted from the census were the Housing Unit Unduplication Operation and the Kill Process.

### 2.4.1  The Housing Unit Unduplication Operation

Prior to the HCUF creation, the Census Bureau developed an automated process that identified duplicate housing units and removed them from the final census counts. The process implemented was conducted in two phases. *Phase 1* identified potential duplicate housing units. Algorithms were developed for identifying MAFIDs that were likely to be duplicates. These algorithms were developed in two approaches. The first approach identified pairs or clusters of MAFIDs that were likely duplicates because the addresses were the same or substantially equivalent. These algorithms were run on the Master Address File (MAF). The second approach identified pairs or clusters of MAFIDs that were likely duplicates because the household (using person data) were the same or substantially equivalent. These algorithms were run on the Decennial Response File 1 (DRF1). For each pair of duplicate MAFIDs, one census ID was retained and the other(s) flagged as a delete on the HCUF Prime address inventory.

*Phase 2* determined the final housing unit status for the potential duplicates. The flagged duplicate MAFIDs were temporarily disregarded in subsequent processing until their final housing unit status was determined. Staff reviewed address information, operational data, and person data to assess the appropriateness of classifying the flagged MAFIDs as duplicate units. MAFIDs were reinstated if they were not likely to represent duplicate housing units, but reflected other situations, such as mover households or instances of questionnaire misdelivery (Nash, 2000).

As a result, there were 2.4 million addresses flagged as potential duplicates for review. A large number of these units were later reinstated (42 percent), however reinstated units were not included in this study.

9

*2.4.2   Kill Process*

As a result of the DMAF building and updating process, addresses were included in the DMAF that did not uniquely identify a housing unit as of April 1, 2000, Census Day. Those addresses made up the "kill universe".  A kill or delete was enforced on addresses if two census field operations confirmed the address did not exist on Census Day.  The kill universe was divided into categories or reasons for the deletes.  For an address to be deletes in the Kill Process, it had to be in at least one of the categories (Treat, 2000).

- A.    Double Delete and no Mail Return (MR)
- B.    Old Delivery Sequence File (DSF) Addresses and no MR
- C.    Undeliverable as Addressed (UAA), Nonresponse Followup (NRFU) delete, not in Coverage Improvement Followup (CIFU), not a CIFU add and no MR
- D.    Update/Leave delete, NRFU delete, not in CIFU, not a CIFU add and no MR
- E.    Urban Update/Leave (UU/L) delete, NRFU delete, not in CIFU, not a CIFU add and no MR
- F.    Update/Enumerate delete and no MR
- G.    NRFU delete, CIFU delete and no MR
- H.    NRFU delete, not in CIFU, not a CIFU add and no MR
- I.    Not in NRFU, CIFU delete and no MR
- J.    Address Field Verification (FV) delete or duplicate, not a NRFU add and not a CIFU add
- K.    Adds from the July 7, 2000 update of the DMAF which were Usual Home Elsewhere (UHE) addresses that were generated from Special Place/Group Quarters which were not allowed to provide a UHE address
- L.    Any MAFID unit with a surviving MAFID

We identified duplicates in the Kill Process using the delete action codes from various census operations.

## 2.5    HCUF Adds

HCUF adds were identified by comparing the A.C.E. extract of the January 2000 DMAF to the HCUF Prime files.  Units that were on the HCUF Prime and not on the January 2000 DMAF address inventory are called HCUF adds in this study.  *Again, the census units reinstated from the Housing Unit Unduplication Operation are not included in this analysis.*

The weighted estimates of the HCUF adds are presented in the results section (Section 4). We used the A.C.E. Final Housing Unit (FHU) match results to determine whether the HCUF adds were correctly added or added in error. We classified adds that were coded as correct enumerations or had unresolved enumeration status as correct adds. We classified adds as in error if they were coded in FHU as erroneous enumerations.

## 2.6 Correctly Added Housing Units

Correctly added HCUF adds refer to the units added to the census inventory that were found by A.C.E. to have existed on Census Day. The following were A.C.E. correct enumeration codes assigned to these cases:

- *Match or M* - The P-sample and E-sample housing units matched in A.C.E.'s Final Housing Unit match operation.

- *Correct enumeration or CE* - The FHU followup interview determined the E-sample housing unit existed as a housing unit on Census Day and was correctly geocoded in the block cluster. The housing unit was not matched to an A.C.E. unit.

- *Geocoded Correctly or GC* - The E-sample housing unit was found in the sample block cluster during the second targeted extended search[6] field followup. It was correctly enumerated in the block cluster.

- *Geocoded in the surrounding block or GS* - The E-sample housing unit was found to be in the surrounding blocks during the A.C.E. second targeted extended search field followup. The E-sample housing unit was counted once and only once in the expanded search area.

- *Possible Matches or P* - The code P was assigned when the E-sample housing unit was a possible match to a P-sample housing unit, but the FHU followup interview was inconclusive or incomplete.

- *Unresolved status or UE, MU, GU* - UE match codes were assigned when there was not enough information on the A.C.E. housing unit followup form to assign a code to the census nonmatched housing unit with certainty. MU match codes were

---

[6]   Targeted Extended Search II was an A.C.E. operation performed to check the geocoding status of erroneously enumerated housing units, housing units added to the DMAF after January 2001 with persons coded as geocoding error, and housing units in List/Enumerate clusters that may have benefitted from an extended search.

assigned when the A.C.E. and census addresses match and there was not enough information on the housing unit followup form to confirm the match as a housing unit with certainty. GU match codes were assigned in targeted extended search (TES) and the E-sample address was not located in the A.C.E. cluster and there was not enough information to determine whether it was in the surrounding block.

## 2.7　Added in Error Housing Units

Added in error HCUF adds consisted of housing units that were added to the census address inventory and were found not to exist on Census Day during the A.C.E. Final Housing Unit operations. The following are erroneous enumeration codes assigned by A.C.E. to these cases:

- *Duplicates or DE* - The E-sample address was found to be duplicated in the census by A.C.E.

- *Not a Housing Unit or EE* - The FHU followup determined the E-sample address was nonresidential or did not exist on Census Day; that is, the address was for a group quarters, business, or the unit was demolished, burned down, uninhabitable or could not be located.

- *Geocoding Error or GE* - The E-sample housing unit existed as a housing unit at the time of the A.C.E. followup interview, but was incorrectly geocoded to the block cluster. As a result, the housing unit is said to be erroneously enumerated in the block cluster.

## 2.8　Original Source of Addresses by Census Operations

Addresses were recorded by various census operations that occurred both before and during the census. An original source variable, which did not exist on the MAF, was defined and created to identify the first operation or file that added the address to the MAF (Vitrano, 2001). Below are the various sources from which addresses were generated.

- **1990 Address Control File (ACF):** These were addresses on file at the Census Bureau in 1990.
- **Address Listing (AL):** A field operation that developed the Census 2000 address list in non-mailout/mailback enumeration areas.
- **Postal Delivery Sequence Files (DSF):** A computerized file containing all delivery point addresses serviced by the U.S. Postal Service (USPS). The files were updated monthly.

- **Block Canvassing (BC):** This operation was a field verification of addresses on the Master Address File as of January 1999.
- **Local Update of Census Addresses (LUCA):** The LUCA programs verified the existence and the residential status of addresses. These updates were attributable to a cooperative effort with local and tribal government.
- **Questionnaire Delivery (QST):** QST refers to any operation where the address list was updated during the initial delivery of a questionnaire or during the actual enumeration. These operations include Update/Leave, Update/Enumerate, List/Enumerate, and remote Alaska.
- **Nonresponse Followup (NRFU):** The NRFU operation was conducted to obtain completed questionnaires from households in the mailback areas that did not respond. In this census operation, enumerators visited addresses from which no census questionnaire was returned by mail.
- **Coverage Improvement Followup (CIFU):** The CIFU was a census field operation in which addresses previously identified as vacant or deleted (nonexistent) from the Master Address File were verified to be sure that their vacant or deleted statuses were correct. Also enumerated in CIFU were new construction cases identified by governmental units, late added addresses identified during Update/Leave and through update partnership efforts with the U.S. Postal Service, and addresses for which mail return questionnaires were lost or returned blank.
- **Be Counted (BC):** The BC program provided a means for people to be included in Census 2000 who believed they were not counted. Unaddressed census questionnaires were placed at selected sites that were easily accessible and frequented by large numbers of people.
- **Telephone Questionnaire Assistance (TQA):** The TQA program was a service provided by telephone centers that answered questions about Census 2000 and census questionnaires.
- **New Construction (NC):** These are address updates from housing units recently built and ready for occupancy or are already occupied.
- **Special Place or Group Quarters (SPGQ):** The address is from the census enumeration of special places and group quarters.
- **Unknown:** These are addresses where type of enumeration area was blank or was known, but no other characteristics about the address were known.
- **Dress Rehearsal:** These are addresses from a census of population and housing conducted in selected areas prior to a decennial census to determine the effectiveness of planned census operations.

# 3. LIMITS

The use of the A.C.E. to evaluate adds and deletes entails certain limitations that must be kept in mind when considering the results of this study.

## 3.1    Timing of January 2000 Decennial Master Address File

The January 2000 DMAF was used in this analysis because this was the version of the DMAF that A.C.E. used for the initial housing unit match.  This choice of date restricts the universe of cases that we examine in our study.  Units deleted from the census inventory before January 1, 2000 are not included in this study, as are units added after January 1, 2000 and subsequently deleted.  Similarly, units added before January 1, 2000 are not included in the evaluation of adds.

## 3.2    The A.C.E. Search Area

The A.C.E. search area refers to the blocks where the A.C.E. searched for matches and for duplicates.  The search area consisted of the A.C.E. block cluster and the first ring of blocks surrounding the block cluster.  The search area poses an important limitation because the A.C.E. could not detect a census duplicate housing unit that was outside the its search area.

This limitation is relevant to the housing units deleted from the census by the Housing Unit Unduplication Operation which were counted as deleted in error.  Some of these could have had duplicated housing units outside the A.C.E. search area.  If the census deleted the correctly geocoded unit while keeping the incorrectly geocoded unit, the study would classify the deletion as incorrect.  Thus the estimates for Housing Unit Unduplication Operation incorrect deletions represent an upper bound.  In contrast, we think this is less of a limitation for housing units deleted by other census operations such as the Kill Process since we counted them as correct deletions if the census identified them as duplicates.

## 3.3    A.C.E. Geocoding Error vs. Census Geocoding Error

There were differences in how the A.C.E. and census defined geocoding error.  The A.C.E. defined a block cluster as a single census collection block or group of geographically contiguous census collection of blocks.  An A.C.E. geocoding error occurred when a unit was located outside of the first ring of blocks surrounding the block cluster.  On the other hand, the census does not list housing units by cluster, but rather by block.  A census geocoding error occurred when the assignment of a housing unit is geocoded to an incorrect collection block.  More analysis on geocoding is found in the

14

report "An Assessment of Addresses on the Master Address File 'Missing' in the Census or Geocoded to the Wrong Collection Block, Evaluation F.15″ (Ruhnke, 2003).

## 3.4     Puerto Rico

The results in this report reflect state-side only estimates.  Estimates for Puerto Rico are not included.

# 4.  RESULTS

## 4.1    What proportion of  DMAF deletes were correctly deleted and erroneously deleted by the census according to their enumeration status in the 2000 Accuracy and Coverage Evaluation?

**"Correctly deleted"** addresses were defined as those units where the A.C.E. and the census were in agreement on the delete status, that is, both the A.C.E. and census determined these units did not exist as housing units on Census Day, or were identified as duplicates of other units (in some cases we counted a census deletion as correct even when the A.C.E. disagreed, as in the case of duplicates identified in the Kill Process). These addresses were classified as erroneous enumerations in the A.C.E. or were identified as duplicates.  Addresses **"deleted in error"** were defined as those deleted housing units that the A.C.E. classified as correct enumerations in the Initial Housing Unit  or where a matching P-sample housing unit was confirmed to exist on Census Day by the Person Interview, and were not determined to be duplicates by the A.C.E. or the Kill Process.

**Table 1a** shows the weighted estimate and percent of addresses that were DMAF deletes as well as the distribution of units correctly deleted or deleted in error by the census according to the A.C.E.  These estimates exclude the DMAF deletes that were reinstated housing units from the Housing Unit Unduplication Operation.

- There was a weighted estimate of approximately 8.5 million DMAF deletes. These are addresses that were on the January 2000 DMAF and not on the HCUF Prime.

- About 85.6 percent of the DMAF deletes were correctly deleted by the census. These were addresses that were coded as erroneous enumerations in A.C.E.  Note that DMAF deletes that were identified as duplicates were counted as correctly deleted. (See Section 2. Methods).

- The difference, about 14.4 percent of the DMAF deletes, which represented a weighted estimate of 1.2 million housing units, did not agree with the enumeration status determined by A.C.E.  These were units that were coded as census nonmatches during the Initial Housing Unit match plus units that were coded as matches in IHU which were coded as  P-sample nonmatches during the Final Housing Unit Match.  These DMAF deletes were correct enumerations in A.C.E. but the census deleted them from the inventory.  See the following Sections 4.1.2 through 4.1.3 for further information on how the "deleted in error" estimates were derived.

**Table 1a.  Status of DMAF Deletes**

| DMAF Deletes | Weighted Count | Percent |
|---|---|---|
| Correctly Deleted | 7,309,409* | 85.6 |
| Deleted in Error | 1,227,343* | 14.4 |
| Total DMAF deletes | 8,536,752** | 100.0 |

*See the following Sections 4.1.1 through 4.1.3, for detailed information on how these estimate were derived.*
***Excludes DMAF deletes that were reinstated housing units from the Housing Unit Unduplication Operation.*

**The following sections and tables provide detailed information on how the "correctly deleted" and "deleted in error" estimates were derived**.

*4.1.1   What were the Initial Housing Unit (IHU) matching results of the DMAF deletes?*

Each unit in the DMAF delete universe was evaluated by A.C.E. in the IHU matching operation, which used computer matching, clerical matching, and a field followup interview to determine the A.C.E. enumeration status.  **Table 1b** shows the matching results from the Initial Housing Unit phase of A.C.E. for all DMAF deletes.

- According to A.C.E., almost half (about 48.8 percent) of the DMAF deletes were "not housing units" on Census Day.  This implies that about half the time A.C.E. and the census were in agreement that the census address was nonresidential or did not exist as a housing unit on Census Day.  These units were determined to be erroneously enumerated by the A.C.E. and thus were correctly deleted by the census.

- About 9.3 percent of the DMAF deletes were identified in IHU as geocoding error.  These housing units existed on Census Day, but were incorrectly geocoded in the A.C.E. sample block.  These units were correctly deleted by the census.

- According to A.C.E., about 6.9 percent of the DMAF deletes were found to be duplicates in the census.  This implies that there were two or more records on the address inventory representing the same housing unit.  Census duplicates found by the A.C.E. were erroneous enumerations and thus were counted as correctly deleted by the census.

17

- About 4.0 percent of the DMAF deletes were census nonmatches and were confirmed to exist as housing units on Census Day during the Housing Unit Followup (HUFU) operation. We examined the census nonmatches further to determine what proportion contributed to the "deleted in error" estimates. See Table 1c in Section 4.1.2 for a discussion of DMAF deletes that were census nonmatches in IHU.

- More than 30 percent of the DMAF deletes were matched in the IHU match operation. These matches were examined further in the Final Housing Unit phase of A.C.E. See Section 4.1.3.

**Table 1b. DMAF Deletes by A.C.E. Initial Housing Unit Matching Results**

| A.C.E. Enumeration Status | Unweighted DMAF Deletes | Weighted DMAF Deletes | Weighted Percent |
|---|---|---|---|
| **Erroneous Enumeration** | | | |
| Not a Housing Unit | 34,384 | 4,166,319 | 48.8 |
| Geocoding Error | 7,641 | 798,811 | 9.4 |
| Census Duplicates | 6,329 | 588,380 | 6.9 |
| Unresolved Units | 316 | 42,640 | 0.5 |
| **Correct Enumeration** | | | |
| Census Nonmatches | 2,728* | 341,280 | 4.0 |
| Matches | 20,651** | 2,599,322 | 30.4 |
| **Total DMAF Deletes** | **72,049** | **8,536,752** | **100.0** |

*Census nonmatches were confirmed to exist as housing units at the time of the A.C.E. housing unit followup interview. See Section 4.1.2, Table 1c below.
** Matches were examined further in FHU phase of A.C.E. see Section 4.1.3, Table 1d.

*4.1.2 What proportion of DMAF deletes coded as census nonmatches during the A.C.E. Initial Housing Unit (IHU)Match were correctly deleted or deleted in error by the census?*

During the A.C.E. IHU matching, census addresses that could not be computer and/or clerically matched were sent to field followup to determine if the address existed as a housing unit at the time of the operation. The HUFU operation was conducted from February 12, 2000 through April 4, 2000.

**Table 1c** shows the unweighted and weighted distribution of the DMAF deletes that were coded as census nonmatches. The row highlighted is the proportion of units that contributed to the "deleted in error" estimate.

- Not all of the census nonmatches contributed to the "deleted in error"estimate. About 16.0 percent of the DMAF deletes that were coded as census nonmatches during IHU were also either surviving MAFIDs of census duplicates, primaries to census duplicates, or census duplicates identified from the Kill Process. As noted in the Methods section, DMAF deletes that were duplicates were counted as correctly deleted in this report.

- However, about 84.1 percent of the DMAF deletes coded by A.C.E. as census nonmatches were deleted in error. These units were confirmed to exist as housing units on Census Day during HUFU and were classified as A.C.E. correct enumerations.

**Table 1c.  DMAF Delete Census Nonmatches in Initial Housing Unit**

|  | Unweighted DMAF deletes | Weighted DMAF deletes | Weighted Percent |
|---|---|---|---|
| Surviving MAFIDs | 28 | 2,807 | 0.8 |
| A.C.E. Primaries | 111 | 14,206 | 4.2 |
| Duplicates in Kill Process | 372 | 37,388 | 11.0 |
| Deleted in Error Census Nonmatches | 2,217* | 286,879 | 84.1 |
| **Total Census Nonmatches** | **2,728** | **341,280** | **100.1\*\*** |

*Confirmed to exist as housing units during A.C.E. HUFU thus deleted in error.*
*\*\*Percent does not add to 100 percent due to rounding error.*

### 4.1.3 *What proportion of DMAF deletes coded as matches in IHU were correctly deleted or deleted in error by the census after the A.C.E. Final Housing Unit (FHU) match?*

The DMAF deletes that were matched during the IHU match phase of A.C.E. processing were further evaluated in the FHU match phase. A.C.E. addresses that were originally matched to DMAF deletes were computer processed as new P-sample nonmatches for FHU processing. If these new P-sample nonmatches could not be clerically matched to a census address, they were coded confirmed to exist on Census Day based on the results of

the person interview or CAPI.  Therefore, the previously matched DMAF deletes are considered "deleted in error".

**Table 1d** shows the unweighted and weighted distribution of the DMAF deletes that matched to a P-sample address during IHU.  It also shows the results of DMAF deletes processed in FHU.  The highlighted row was the proportion of units that contributed to the "deleted in error" estimate.

- Matched records represented 30.4 percent (2.6 million weighted housing units) of the DMAF deletes found by A.C.E. in IHU.

- Over half of the unweighted P-sample addresses that were matched to DMAF deletes in IHU were subsampled out due to large block subsampling.  **These units carry a weight of zero, thus having no affect on the estimates**.

- About 1.2 million of the matched records were identified to be correct deletions because they were either CAPI removed, Surviving MAFIDs, primaries or census identified duplicates.

  - About 17.2 percent of the DMAF deletes that matched P-sample addresses were deleted during the A.C.E. CAPI.  These units were correctly deleted by the census.

  - About 27.4 percent of the DMAF deletes that matched a P-sample address in IHU were either surviving MAFIDs of census duplicates, A.C.E. primaries to duplicates, or census duplicates identified from the Kill Process.  These units were counted as correctly deleted.

- The remainder, about 1.4 million DMAF deletes that were matched to A.C.E. addresses in IHU (and were not CAPI removed or census duplicates) were evaluated based on the FHU. Below are the results:

  - As with census nonmatches, P-sample nonmatches (that originally matched a DMAF delete) which were confirmed to exist as housing units on Census Day during person interviewing were classified as existing as housing units on Census Day in A.C.E. These matching DMAF deletes (if they were not identified as duplicates) contributed to the "deleted in error" weighted estimate of 940,464 addresses, 36.2 percent of the total DMAF deletes that matched in Initial Housing Unit Operations.

  - A weighted estimate of 475,081 (18.3 percent) A.C.E. addresses that matched DMAF deletes in IHU were matched to other census addresses during FHU processing. These DMAF deletes were counted as correctly deleted by the census.

  - About one percent of the A.C.E. addresses that matched DMAF deletes were removed from the P sample because they were duplicates of another A.C.E. address, the address was geocoded to a an incorrect block, the address was unresolved in FHU, or the address was incorrectly included in the P sample. Since these addresses were removed from the P sample, the matching DMAF delete addresses were counted as correctly deleted by the census.

**Table 1d. Status of DMAF Deletes Matched in Initial Housing Unit**

| | Unweighted DMAF Deletes | Weighted DMAF Deletes | Weighted Percent |
|---|---|---|---|
| CAPI Removed | 1,492 | 447,689 | 17.2 |
| Primaries | 1,213 | 334,093 | 12.9 |
| Surviving MAFIDs | 117 | 42,530 | 1.6 |
| Duplicates in the Kill Process | 1,044 | 334,175 | 12.9 |
| Subsampled out in A.C.E. Large Block Subsampling | 12,588 | 0 | 0.0 |
| **Processed in Final Housing Unit** | **4,197** | **1,440,836** | **55.4** |
| Deleted in error P-sample Nonmatches | 2,756 | 940,464** | 36.2 |
| Matches | 1,363 | 475,081 | 18.3 |
| Removed from P-sample | 78 | 25,291 | 1.0 |
| **Total DMAF Delete Matches** | **20,651** | **2,599,322** | **100.1*** |

*Percent does not add up to 100 percent due to rounding error.
** Confirmed to exist as housing units during A.C.E. Person Interview, thus deleted in error.

*4.1.4   How many correctly deleted DMAF deletes were there by reason for delete?*

**Table 1e** presents the correctly deleted DMAF deletes by reason for delete as determined by A.C.E.  The table shows the reasons for delete by A.C.E. and by census.  There were about 7.3 million weighted DMAF deletes classified as erroneous enumerations in A.C.E. or were duplicates in the census.

•      Of the correctly deleted census housing units, 63.1 percent were not a housing unit or did not exist on Census Day according to the Accuracy and Coverage Evaluation.  This includes the 57.0 percent identified as not a housing unit plus the 6.1 percent CAPI removed.

•      The A.C.E.  duplicate addresses accounted for a total of 12.8 percent of the correctly deleted Decennial Master Address File deletes.  This includes the 8.0

percent coded as duplicate enumerations by A.C.E. plus another 4.8 percent that were coded primaries by the A.C.E.

- About 10.9 percent of the DMAF deletes that were correctly deleted were coded as geocoding error in A.C.E. These units existed at the time of the A.C.E. followup interview, but were outside the A.C.E. sample block cluster, thus incorrectly included in the A.C.E. block.

**Table 1e. DMAF Deletes Correctly Deleted by Reason for Delete**

| Reason for delete | Weighted DMAF Deletes | Weighted Percent |
|---|---|---|
| A.C.E. Not a Housing Unit | 4,166,319 | 57.0 |
| A.C.E. Geocoding Error | 798,811 | 10.9 |
| A.C.E. Duplicates | 588,380 | 8.0 |
| Matches in FHU | 475,081 | 6.5 |
| A.C.E. identified in CAPI as not a housing unit | 447,689 | 6.1 |
| Identified as Duplicates in the Kill Process | 371,562 | 5.1 |
| A.C.E. identified as duplicates (primaries) | 348,299 | 4.8 |
| Census Surviving MAFIDs | 45,336 | 0.6 |
| A.C.E. Unresolved housing unit status | 42,640 | 0.6 |
| Removed from P-sample in FHU | 25,291 | 0.3 |
| **Total** | **7,309,409** | **99.9*** |

*Percent does not add to 100 percent due to rounding error.
**The totals assigned for the categories represent the subset of that category that does not overlap with the higher priority categories. See Methods Section 2.2 for more details.*

**4.2 Which census operations contributed to the DMAF deletes?**

As previously stated (Methods Section 2.4), there were four operations that identified housing units on the DMAF that were excluded from the census.  We examine two of the census delete processes in Tables 2a and 2b.  Tables 2a and 2b show the estimate of DMAF deletes that were deleted by the Housing Unit Unduplication Operation and the Kill Process.  These two operations accounted for 97.4 percent (999,850 housing units and 7,317,697 housing units, respectively) of the DMAF deletes in this study.

*4.2.1  Housing Unit Unduplication Operation*

**Table 2a** presents the DMAF deletes removed from the inventory as a result of the Housing Unit Unduplication Operation.  The unduplication process accounted for about 12 percent (999,850 housing units) of the total DMAF deletes (8.5 million housing units).

- Of the approximately 1.4 million housing units deleted in the Housing Unit Unduplication Operation, about 72.9 percent of the units were DMAF deletes in A.C.E.

- According to A.C.E., about 36.0 percent (359,966 housing units) of the total DMAF deletes deleted in the Housing Unit Unduplication Operation were deleted erroneously.  This operation accounted for 29.3 percent of the total weighted number of deletes in error (1,227,344 housing units).  (Note this estimate represents an upper bound because the A.C.E. did not detect duplicates beyond its search area; see Section 3.2, Limits.)

- About 64.0 percent of the DMAF deletes deleted in the Housing Unit Unduplication Operation were correctly deleted.

**Table 2a. DMAF Deletes in the Housing Unit Unduplication Operation** (%)

| Census Operation | DMAF deletes Correctly Deleted | DMAF deletes Deleted in error | Number of DMAF deletes | Total # of HUs deleted by the Housing Unit Unduplication Operation |
|---|---|---|---|---|
| Housing Unit Unduplication | | | | |
| Weighted | 639,884 (64.0) | 359,966 (36.0) | 999,850 (100.0) | 1,371,320* |

*4.2.2   Kill Process*

As discussed in Methods Section 2.4.2, the Kill Process was made up of addresses that did not uniquely identify housing units on Census Day.  Of the DMAF deletes identified by A.C.E. (8.5 million housing units), most of them (85.7 percent or 7.3 million housing units) were removed from the census by the Kill Process.  A total of about 9.0 million housing units were deleted by the Kill Process.

**Table 2b** presents the estimates of the DMAF deletes deleted by the Kill Process that were correctly deleted and deleted in error according to A.C.E.

- We estimate that the Kill Process deleted in error 752,705 housing units from the census.  This accounts for 61.3 percent of the total weighted number of deletes in error (1,227,344); however, this estimate only represents about 10.3 percent of the DMAF deletes that were deleted by the Kill Process.

- According to A.C.E., about 90 percent of the DMAF deletes deleted by the Kill Process were correctly deleted by the census.

**Table 2b. DMAF Deletes in the Kill Process** (%)

| Census Operation | DMAF deletes Correctly Deleted | DMAF deletes Deleted in error | Number of DMAF deletes | Total # of HUs deleted by the Kill Process |
|---|---|---|---|---|
| Kill Process | | | | |
| DMAF deletes | 6,564,992 (89.7) | 752,705 (10.3) | 7,317,697 (100.0) | 8,977,958* |

*\* Count represents stateside only.*

## 4.3   What were the reasons units were deleted by the Kill Process?

As shown in the previous table (Table 2b), of the units that were excluded from the census by the Kill Process (7.3 million housing units), about 10 percent or 752,705 housing units were deleted in error according to A.C.E.  In the Kill Process, deleted units were defined by categories in which an address could be classified as an invalid housing unit.  For an address to be deleted, it had to satisfy at least one of the defined categories.  In each category, at least two census operations confirmed the address as a delete.  See Section 2.4.2 for categories in the Kill Process.

In Tables 3a and 3b, we look at the units that were deleted in error by the Kill Process and the reasons documented by the Kill Process for deleting the units.

The goal of the Kill Process was to identify addresses that were not housing units or were duplicates to remove them from the census. One example of the type of unit that was excluded from the census as a result of this process was:

- there was no census form returned for the unit,
- the unit was deleted in NRFU, and
- the unit was confirmed as a delete in the CIFU operation.

*DMAF deletes deleted in error by the Kill Process*

**Table 3a** provides the distribution of kills for the deleted in error DMAF deletes by kill reason.

- The kill reason with the largest number of deleted in error cases was the situation where addresses were Undeliverable as Addressed by the US Postal Service and were deleted in the NRFU operation. These cases represented about 35.2 percent of the DMAF deletes deleted in error by the Kill Process.

- About 23.1 percent of the deletes in error removed by the Kill Process were deleted because there was no mail return and both the NRFU and CIFU operations deleted the unit.

- About 10.3 percent of the addresses deleted in error by the Kill Process were a result of Old Delivery Sequence File (DSF) addresses and no mail return. These addresses were not in the November 1999, February 2000, and April 2000 DSFs.

- About 8.9 percent of the deleted in error units were deleted because there was no mail return and NRFU deleted the unit. These units were not in the CIFU workload and were not added during CIFU processing.

26

**Table 3a. Addresses Erroneously Deleted in the Kill Process by Reason for Kill** (column %)

| Reason for Kill | Weighted DMAF deletes |
|---|---|
| BC Delete & LUCA 98 FV Delete or BC could not locate, LUCA 98 add, and LUCA 98 FV delete; no mail return (A) | 48,288 (6.4) |
| Old Delivery Sequence File delete and no mail return (B) | 77,763 (10.3) |
| UAA & NRFU Delete (C) | 265,063 (35.2) |
| UU/L or U/L Delete, NRFU Delete, not in CIFU, not a CIFU add; no mail return (D/E) | 77,613 (10.3) |
| U/E Delete & no mail return (F) | 3,127 (0.4) |
| NRFU Delete, CIFU Delete; no mail return (G) | 173,515 (23.1) |
| NRFU Delete, not in CIFU, not a CIFU add; no mail return (H) | 67,163 (8.9) |
| Not in NRFU, CIFU Delete; no mail return (I) | 20,857 (2.8) |
| Address FV Delete or Duplicate, not a NRFU add / CIFU add (J) | 19,316 (2.6) |
| Total DMAF deletes | 752,705 (100.0) |

### 4.3.2   *DMAF deletes deleted in error by NRFU*

For each category in the Kill Process, the census operation that deleted housing units applied an action code documenting the reason(s) for the delete.  The kill categories where NRFU was one of the reasons were:

> C - UAA & NRFU delete,
> D/E - UU/L or U/L delete, NRFU delete, not in CIFU, not a CIFU add and no mail return,
> G - NRFU delete, CIFU delete, and no mail return, and
> H - NRFU delete, not in CIFU, not a CIFU add, and no mail return.

**Table 3b** presents the DMAF deletes deleted in error by NRFU delete categories.

- Category C (UAA & NRFU delete) represents the largest number of DMAF deletes in error from the Kill Process (35.2 percent, from Table 3a).  NRFU deleted more than 40 percent of the addresses because the units were found to be either open to the elements, condemned, or under construction.  For about 32.7 percent of the units, the NRFU enumerator was unable to locate the unit.

- As shown in Table 3a, Category D/E accounted for 10.3 percent of DMAF deletes deleted in error by the Kill Process.  About 37.8 percent of these units were found in NRFU to be either open to the elements, condemned, or under construction.  More than 34.2 percent of the units NRFU could not locate and about 16.1 percent of the housing units NRFU found to be demolished.

- Similar findings in the "Cannot Locate" and "Open to elements, condemned, or under construction" reasons for deleting the unit were found for both categories G and H.  For about 33 percent of the units (33.4 percent and 32.9 percent, respectively), NRFU was unable to locate the unit.  More than 30 percent of the units (33.6 percent in category G and 30.6 percent in category H) were found to be open to the elements, condemned, or under construction by NRFU.

- For Category G, NRFU found about 27.0 percent of those units to be nonresidential and about 6.0 percent were found to be demolished.

**Table 3b. Kill Universe Categories C, D/E, G, and H: Addresses Deleted in Error by Nonresponse Followup** (column %)

| | **NRFU reason for delete** | | | | |
|---|---|---|---|---|---|
| Kill Process reasons for delete | Demolished | Cannot Locate | Nonresidential | Other (open to elements, condemned, under construction) | Total |
| UAA & NRFU Delete (C) | 29,069 (11.0) | 86,590 (32.7) | 41,666 (15.7) | 107,738 (40.6) | 265,063 (100.0) |
| UU/L or U/L Delete, NRFU Delete, not in CIFU, not a CIFU add; no mail return (D/E) | 12,468 (16.1) | 26,557 (34.2) | 9,248 (11.9) | 29,340 (37.8) | 77,613 (100.0) |
| NRFU Delete, CIFU Delete, no mail return (G) | 10,433 (6.0) | 57,902 (33.4) | 46,916 (27.0) | 58,264 (33.6) | 173,515 (100.0) |
| NRFU Delete, not in CIFU, not a CIFU add; no mail return (H) | 4,330 (6.4) | 22,110 (32.9) | 20,138 (30.0) | 20,585 (30.6) | 67,163 (100.0) |

### 4.4 Were the housing units identified as census HCUF adds correct or erroneous enumerations in A.C.E.?

After the IHU match operation of A.C.E., housing units were both added and deleted to the census list.  In this section, we look at those units that were added to the inventory. We used the A.C.E. Final Housing Unit (FHU) match results to evaluate the HCUF add housing units.

**Table 4** shows the weighted estimate of HCUF adds.

- According to A.C.E., there was an total weighted estimate of approximately 3.9 million HCUF adds in FHU.

- About 83.9 percent of the HCUF adds were correctly added to the census housing unit inventory.  These units were found to exist as housing units on Census Day and no evidence of duplication was found by A.C.E.

- The remainder, about 16.1 percent of the HCUF adds were found to be erroneous enumerations by A.C.E., thus added in error.  These cases included units that were found to be duplicates by A.C.E., not a housing unit on Census Day, or geocoded to an incorrect block.

#### Table 4. Status of HCUF Adds (Weighted)

| HCUF Adds | Weighted Total HCUF Adds (%) |
|---|---|
| Correctly Added | 3,235,099 (83.9) |
| Added in Error | 622,282 (16.1) |
| Total HCUF Adds | 3,857,381 (100.0) |

### 4.5    What were the Final Housing Unit matching results of the HCUF adds?

**Table 5a** presents the distribution of HCUF adds by the A.C.E. FHU matching results.

- Of the total HCUF adds, 83.9 percent were classified by A.C.E. as correct enumerations; that is, we determined they were correctly added.

    - Of the units added to the inventory, about 46.0 percent of the HCUF adds were matched in FHU. A.C.E. confirmed the unit to exist as a housing unit on Census Day.

    - About 36.1 percent of the HCUF adds were census nonmatches in FHU. These units were confirmed to exist as housing units on Census Day by A.C.E. and were deemed correctly added units by the census.

- The remainder of the HCUF adds (16.1 percent) were classified by A.C.E. as erroneous enumeration, and therefore we classified them as added in error.

    - According to A.C.E., about 8.1 percent of the HCUF adds were "not housing units" on Census Day. This implies that census added units where the census address was nonresidential or did not exist on Census Day. These units were added to the inventory erroneously.

    - About 4.2 percent were found to be geocoding error in FHU. These housing units existed on Census Day, but were incorrectly geocoded in the A.C.E. block.

    - About 3.8 percent of the HCUF adds were found to be duplicates of other units in the inventory.

**Table 5a. Weighted HCUF Adds Added in Error by the Census**

| DMAF Delete A.C.E. Housing Unit Status | Unweighted HCUF Adds | Weighted HCUF Adds |
|---|---|---|
| **Correct Enumerations** | **25,294** | **3,235,099 (83.9)** |
| Matches | 20,561 | 1,775,631 (46.0) |
| Possible matches | 13 | 815 (0.0) |
| Census nonmatches | 4,523 | 1,394,332 (36.1) |
| Unresolved Housing Unit Status | 197 | 64,321 (1.7) |
| **Erroneous Enumerations** | **2,749** | **622,282 (16.1)** |
| Not a Housing Unit | 1,267 | 314,019 (8.1) |
| Geocoding Error | 1,001 | 163,581 (4.2) |
| Census Duplicates | 481 | 144,682 (3.8) |
| **Total** | 28,043 | 3,857,381 (100.0) |

*4.5.1  What were the HCUF adds added in error by type of erroneous enumeration status?*

**Table 5b** presents the distribution of HCUF adds added in error by type of erroneous enumeration status.  The erroneous enumeration codes in A.C.E. were not a housing unit, census duplicates, or  geocoding error.  According to A.C.E., there were an estimated 622,282 weighted HCUF addresses added in error by the census.

- According to A.C.E., over half of the HCUF adds erroneously added were not housing units on Census Day (50.5 percent).

- Geocoding errors accounted for 26.3 percent and census duplicates accounted for 23.3 percent of the HCUF adds erroneously added.

**Table 5b.  Added in Error HCUF Adds by Type of Erroneous Enumeration in A.C.E.**

| Type of Erroneous Enumeration | Unweighted Total HCUF Adds | Total HCUF Addresses Added in Error |
|---|---|---|
| Not a Housing Unit | 1,267 | 314,019 (50.5) |
| Geocoding Error | 1,001 | 163,581 (26.3) |
| Census Duplicates | 481 | 144,682 (23.3) |
| **Total** | **2,749** | **622,282 (100.1)*** |

*Percent does not add to 100 percent due to rounding error.*

### 4.6    What were the HCUF Adds estimates by type of structure?

**Table 6** gives the weighted estimates of HCUF adds by type of structure.

- Overall, there was a much larger proportion of single units (72.4 percent) than small multi-units, 2-9 units (12.9 percent) and large multi-units, 10 or more units (14.7 percent) added to the address inventory.

- Similarly, most of the units correctly added (72.9 percent), as well as added in error (69.7 percent) were single units.

- Of the total HCUF adds added in error by type of structure, single units had the largest proportion (69.7 percent).  Small multi-units followed with 20.8 percent added in error while large multi-units had only about 9.5 percent of adds in error.

33

**Table 6. Weighted HCUF Adds by Type of Structure** (%)*

| Type of Structure (Number of Units) | Correctly Added | Added in Error | Total HCUF Adds |
|---|---|---|---|
| Single units (1) | 2,358,806 (72.9) | 433,676 (69.7) | 2,792,482 (72.4) |
| Small Multi-units (2-9 ) | 368,857 (11.4) | 129,370 (20.8) | 498,227 (12.9) |
| Large Multi-units (10+) | 507,436 (15.7) | 59,236 (9.5) | 566,672 (14.7) |
| Total | 3,235,099 (100.0) | 622,282 (100.0) | 3,857,381 (100.0) |

*Column percent*

### 4.7    What were the HCUF adds by census region?

**Table 7** presents the distribution of HCUF adds by census region.

- According to A.C.E., more than eighty percent of the census adds for each region were correctly added by the census.

- The South region had the largest proportion of census adds (45.9 percent). According to A.C.E., the South region also had the largest proportion of adds correctly added (46.7 percent) and added in error (41.9 percent).

**Table 7.  Weighted HCUF Adds by Census Region** (%)*

| | Northeast | Midwest | South | West | Total |
|---|---|---|---|---|---|
| Census Correctly Added | 556,983 (17.2) (82.4) | 513,788 (15.9) (84.5) | 1,509,895 (46.7) (85.3) | 654,433 (20.2) (81.5) | 3,235,099 (100.0) (83.9) |
| Census Added in Error | 119,088 (19.1) (17.6) | 93,909 (15.1) (15.5) | 260,441 (41.9) (14.7) | 148,844 (23.9) (18.5) | 622,282 (100.0) (16.1) |
| Total | 676,070 (17.5) (100.0) | 607,697 (15.8) (100.0) | 1,770,337 (45.9) (100.0) | 803,277 (20.8) (100.0) | 3,857,381 (100.0) (100.0) |

*The first percent compares correct or incorrect adds by census region.  The second percent is within census region.*

**4.8    What was the original source on the Master Address File for the HCUF adds?**

Evaluations of the MAF building operations required identification of the source of every address on the MAF.  An original source variable was defined that identified the first operation to add the address to the MAF.  **Table 8** presents the distribution of weighted estimates and percent of HCUF adds by the original source and whether the adds were correctly added or added in error as determined by A.C.E.

- Among the Master Address File original sources, Questionnaire Delivery furnished the most (over 2 million or 53.3 percent Hundred Percent Census Unedited file adds to the census inventory).  Although the Accuracy and Coverage Evaluation confirmed most of these adds were correctly added (56.5 percent), Questionnaire Delivery also added the most units in error  (36.6 percent).  Questionnaire Delivery[7] included the following census operations: Update/Leave, Update/Enumerate, Urban Update/Leave, List/Enumerate, and Remote Alaska.

- The Delivery Sequence File (DSF) addresses accounted for 23.9 percent of the HCUF adds; according to A.C.E., 87.4 percent of the DSF addresses existed on Census Day and therefore were correctly added to the address inventory.

- NRFU accounted for 7.1 percent of the HCUF adds, and A.C.E. confirmed that 66.2 percent of the NRFU HCUF adds were correctly added to the census address inventory.  This operation also accounted for 33.8 percent of adds in error.

- Addresses on the 1990 ACF remain valuable.  HCUF add addresses from the 1990 ACF accounted for 5.3 percent of the total HCUF adds, and A.C.E. confirmed that 70.9 percent of the 1990 ACF addresses were correctly added by the census.

- New construction accounted for about 2.3 percent of the HCUF adds and according to A.C.E. 74.6 percent of these were correctly added.

- Block Canvassing (0.9 percent) and Address Listing (0.5 percent) each accounted for less than one percent of the HCUF adds.  According to A.C.E., about 95.9 percent of the HCUF adds were correctly added in Block Canvassing and about 69.3 percent of the HCUF adds were correctly added by Address Listing.

---

[7]    The counts and percentages for Questionnaire Delivery by type of enumeration areas are presented in Appendix, Table A1.  Also shown in Appendix A are the weighted estimates of HCUF adds by TEA.

**Table 8. Weighted HCUF Adds by Original Source** (%)*

| Source of the HCUF Adds | HCUF Adds Correctly Added by Census | HCUF Adds Added in Error by Census | Total HCUF Adds by Source |
|---|---|---|---|
| 1990 Address Control File (ACF) | 146,087 (4.5) (70.9) | 59,935 (9.6) (29.1) | 206,022 (5.3) (100.0) |
| Local Update of Census Addresses (LUCA) | 39,985 (1.2) (65.4) | 21,143 (3.4) (34.6) | 61,128 (1.6) (100.00) |
| Block Canvassing | 31,542 (1.0) (95.9) | 1,342 (0.22) (4.1) | 32,883 (0.9) (100.00) |
| Delivery Sequence File (DSF) | 804,757 (24.9) (87.4) | 115,875 (18.6) (12.6) | 920,632 (23.9) (100.00) |
| Address Listing | 13,736 (0.4) (69.3) | 6,082 (1.0) (30.7) | 19,818 (0.5) (100.00) |
| Questionnaire Delivery (QST) | 1,829,049 (56.5) (88.9) | 227,703 (36.6) (11.1) | 2,056,751 (53.3) (100.00) |
| New Construction (NC) | 64,643 (2.0) (74.6) | 22,032 (3.5) (25.4) | 86,675 (2.3) (100.00) |
| Nonresponse Followup (NRFU) | 180,156 (5.6) (66.2) | 92,061 (14.8) (33.8) | 272,217 (7.1) (100.00) |
| Coverage Improvement Followup (CIFU) | 42,017 (1.3) (65.9) | 21,746 (3.5) (34.1) | 63,762 (1.7) (100.00) |
| Telephone Questionnaire Assistance/ Be Counted  (TQA/BC) | 70,714 (2.2) (73.7) | 25,251 (4.1) (26.3) | 95,965 (2.5) (100.00) |
| Dress Rehearsal / Special Place/Group Quarters | 400 (0.0) (43.9) | 512 (0.1) (56.1) | 912 (0.0) (100.00) |
| Unknown - TEAs 1 - 9 | 12,016 (0.4) (29.6) | 28,600 (4.6) (70.4) | 40,615 (1.1) (100.00) |
| Total | 3,235,099 (100.0) (83.9) | 622,282 (100.0) (16.1) | 3,857,381 (100.00) (100.00) |

* *Column percent; row percent*

**4.9    What was the original source of the HCUF adds that were classified as duplicates by A.C.E.?**

**Table 9** gives the weighted estimates and percent of HCUF adds that were duplicates in A.C.E. by original source. As previously shown in Table 5a, duplicates accounted for 3.8 percent of the HCUF adds. Addresses coded as duplicates were erroneous enumerations in A.C.E. and therefore are classified as added in error to the census inventory of addresses.

- Questionnaire Delivery accounted for 41.8 percent of the HCUF adds that were duplicates.

- The Delivery Sequence File and Nonresponse Followup MAF building sources accounted for the next two highest proportions of duplicate addresses added in error (20.6 percent vs. 20.5 percent, respectively).

**Table 9. Weighted HCUF Adds Classified as Duplicates by Original Source** (percent)

| Source of the Duplicate HCUF Adds | Duplicate HCUF Adds |
|---|---:|
| 1990 Address Control File (ACF) | 2,661 (1.8) |
| Local Update of Census Addresses (LUCA) | 3,372 (2.3) |
| Delivery Sequence File (DSF) | 29,739 (20.6) |
| Address Listing | 705 (0.5) |
| Questionnaire Delivery (QST) | 60,471 (41.8) |
| New Construction (NC) | 1,996 (1.4) |
| Nonresponse Followup (NRFU) | 29,701 (20.5) |
| Coverage Improvement Followup (CIFU) | 4,680 (3.2) |
| Telephone Questionnaire Assistance / Be Counted | 9,322 (6.4) |
| Dress Rehearsal / Special Place/Group Quarters | 236 (0.2) |
| Unknown - TEAs 1 - 9 | 1,799 (1.2) |
| Total | 144,682 (100.0) |

# 5. CONCLUSIONS AND RECOMMENDATIONS

The objective of this report was to use results from the A.C.E. to evaluate the changes in the Census 2000 housing unit inventory between the creation of the January 2000 DMAF and the HCUF. The analysis utilizes results from the A.C.E. housing unit operations that were conducted in two parts, the Initial Housing Unit and Final Housing Unit phases. Still, as discussed earlier, the evaluation entails some important limitations. In particular, the arbitrary nature of January 1, 2000 as the reference date for defining adds and deletes, the exclusion of the reinstated housing units from the study, and the inability of the A.C.E. to detect duplicates outside its search area.

**The major conclusions of this study follow:**

**Most (85.6 percent) of the housing units census deleted were correctly deleted.** That is, the A.C.E. was in agreement with the census that the address did not exist as a housing unit or was a duplicate on Census Day. Of the correctly deleted census housing units, 63.1 percent were not a housing unit or did not exist on Census Day according to the A.C.E. A.C.E. duplicate addresses accounted for a total of 12.8 percent of the correctly deleted DMAF deletes.

**About 14.4 percent of the DMAF deleted units were deleted in error according to the A.C.E.** This represents a weighted estimate of 1.2 million housing units which may have been deleted in error by the census. The A.C.E. confirmed the housing units to exist on Census Day and determined there was no evidence of duplication. These units were visited by the A.C.E. interviewers during the Initial Housing Unit Followup operation (which took place between February 12, 2000 and April 4, 2000) or the Person Interview (which took place from June 19, 2000 through August 18, 2000) to determine if the addresses existed as a housing units on Census Day.

These DMAF deletes that the census deleted in error contributed substantially to census omissions. The Census 2000 Housing Unit Coverage Study, Evaluation O.3, found the national estimate of omissions was 3.62 percent or about four million P-sample nonmatches (Barrett et al, 2003).

**Of the census operations that excluded units from the census, the Kill Process (61.3 percent) and the Housing Unit Unduplication Operation (29.3 percent) accounted for 90.7 percent of the 1.2 million erroneous deletes. Even so, the Kill Process deleted units correctly most of the time (89.7 percent)and the Housing Unit Unduplication Operation correctly deleted units more than half of the time (64.0 percent).** Note that the estimate for the erroneously deleted units by the Housing Unit Unduplication Operation is an upper bound because of the limitations of the A.C.E. search area. Some of these units could have been duplicated housing units outside the

search area.  If the census deleted the correctly geocoded unit, which was inside the search area, while keeping the incorrectly geocoded unit then our study classified the unit as erroneously deleted.

**Most (83.9 percent) of the housing units Census 2000 added were found to be correct enumerations by the A.C.E.**  Results from the A.C.E. Final Housing Unit phase found that these units existed as housing units on Census Day and were neither duplicates nor geocoding error.

**About 16.1 percent of the HCUF adds were added in error by the census.**  The Final Housing Unit Phase of A.C.E. classified these units as erroneous enumerations; that is, they either were not a housing unit on Census Day, were duplicates of another address, or were geocoded to the incorrect A.C.E. cluster.

### Recommendations

The insights this evaluation yielded and the questions it raised have led to the following recommendations:

- Conduct research on the census Kill Process so that it deletes fewer valid housing units.  Its error rate was small as a percentage, but there were enough incorrect deletions to make efforts at reduction important.  This evaluation found that the Kill Process deleted many of the housing units that should have been included in the census because addresses were undeliverable as addressed by the United States Postal Service and were deleted from the address inventory during Nonresponse Followup.  For an address to be deleted, at least two census operations had to confirm the address as a delete.  We recommend further research be conducted that evaluates the use of the "undeliverable as addressed" addresses as one of the determining factors for a kill.

- Conduct a housing unit coverage study for the 2010 Census but design it with more emphasis on evaluating census processes.

- Obtain better estimates of gross census error as opposed to concentrating on net coverage error.  For example, for the housing unit followup, have a consistently applied and wide search area.  This will allow us to better identify census geocoding error and distinguish geocoding error from the housing units that do not exist, and also to identify housing unit duplication that is beyond the A.C.E. search area as defined for 2000.

- Conduct research to refine procedures for identifying and removing units that are believed to be duplicate housing units. This evaluation found that the Housing Unit Unduplication Operation may have deleted many units which should have been included in the census. Planning and work has already begun on building a housing unit unduplication process into the 2010 Census.

- Conduct further research that examines the breakdown of housing unit status by census field operations, with emphasis on status obtained in NRFU and CIFU Operations.

# 6. REFERENCES

Alberti, Nick (2001) "Specifications for Assigning the Housing Unit Status and Population Count on the Hundred Percent Unedited File Prior to Imputation for Unclassified Units," *DSSD Census 2000 Procedures and Operations Memorandum Series, D-14,* January 19, 2001.

Barrett, Diane, Beaghen, Michael, Smith, Damon and Burcham, Joseph (2003) *"Census 2000 Housing Unit Coverage Study, Evaluation O.3,"* February 21, 2003.

Barrett, Diane, Byrne, Rosemary, Adams Tamara (2002) "Operational Plan for Accuracy and Coverage Evaluation for Census 2000," *DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-TL-06,* April 30, 2002.

Beaghen, Michael (2001) "Specifications for the Computer Processing for Final Housing Unit Match," *DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-FH-01,* May 11, 2001.

Chen, Inez and Kilmer, Andrew (2002) "Census 2000: Overview of Count Imputation - Reissue of Q-2," *DSSD Census 2000 Procedures and Operations Memorandum Series, Q-78,* March 18, 2002.

Childers, Danny (2001) "Accuracy and Coverage Evaluation: The Design Document," *DSSD Census 2000 Procedures and Operations Memorandum Series Chapter, S-DT-1,* January 26, 2001.

Cromar, Ryan and Farber, James (2000) "Accuracy and Coverage Evaluation: Final Large Block Cluster Subsampling Specifications," DSSD Census 2000 Procedures and Operations Memorandum Series, R-38, December 7, 2000.

Jones, Charisse (2001) "Specifications for Extracting Data from the Hundred Percent Census Unedited File for the Census 2000 Accuracy and Coverage Evaluation," *DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-CO-04,* February 12, 2001.

Jones, Charisse (1999) "Specifications for Extracting Data from the Decennial Master Address File for the Census 2000 Accuracy and Coverage Evaluation," *DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-CO-02,* November 5, 1999.

Nash, Fay F. (2000) "Overview of the Duplicate Housing Unit Operations," *Census 2000 Informational Memorandum No.78,* November 7, 2000.

Philipp, Dan (2002) "2000 Decennial Census Documentation: Hundred percent Census Unedited File (HCUF)," *2000 Decennial Census HCUF Documentation Final (version 3),* April 11, 2002.

Ruhnke, Megan C. (2003) "Final Report for An Assessment of Addresses on the Master Address File 'Missing' in the Census or Geocoded to the Wrong Collection Block, Evaluation F.15, August 22, 2003.

Treat, James B. (2000) "Specification of the Kill Universe on the Decennial Master Address File for Census 2000," *DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter D-13*, December 21, 2000.

Vitrano, Frank (2001) "Determining Original Source for the November 2000 Master Address File for Evaluation Purposes," *Planning, Research, and Evaluation Division TXE/2010 Memorandum Series: MAF-EXT-S-01*, March 5, 2001.

ZuWallack, Randy (1999) "Accuracy and Coverage Evaluation (A.C.E.) Survey: Sample Summary File and Sample Design File Documentation," *DSSD Census 2000 Procedures and Operations Memorandum Series R-4*, March 30, 1999.

# 7. APPENDIX

**Table A1** presents the Questionnaire Delivery adds, correctly added and added in error, by type of enumeration areas.

**Table A1. Questionnaire Delivery Adds by TEA**

| Type of Enumeration Area | Correctly Added | Added in Error | Total HCUF Adds |
|---|---|---|---|
| Mailout/Mailback | 15,879 | 22,757 | 38,636 |
| Update/Leave | 1,322,192 | 140,520 | 1,462,712 |
| List/Enumerate | 358,512 | 39,463 | 397,975 |
| Rural Update/Enumerate | 82,567 | 20,016 | 102,583 |
| Urban Update/Leave | 1,825 | 851 | 2,676 |
| Urban Update/Enumerate | 1,886 | 54 | 1,940 |
| Update/Leave from Mailout/Mailback areas | 46,188 | 4,041 | 50,229 |
| Total | 1,829,049 | 227,702 | 2,056,751 |

Table A2 presents the HCUF adds, correctly added and added in error, by type of enumeration areas.

**Table A2. Hundred Percent Census Unedited File (HCUF) Adds by TEA** (%)

| Type of Enumeration Area | Correctly Added | Added in Error | Total |
|---|---|---|---|
| Mailout/Mailback | 1,286,237 (39.8) | 342,087 (55.0) | 1,628,324 (42.2) |
| Update/Leave | 1,444,815 (44.7) | 211,243 (34.0) | 1,656,059 (42.9) |
| List/Enumerate | 362,609 (11.2) | 41,616 (6.7) | 404,225 (10.5) |
| Rural Update/Enumerate | 82,778 (2.6) | 20,150 (3.2) | 102,928 (2.7) |
| Urban Update/Leave | 8,255 (0.3) | 1,402 (0.2) | 9,657 (0.3) |
| Urban Update/Enumerate | 2,048 (0.1) | 54 (0.01) | 2,103 (0.1) |
| Update/Leave from Mailout/Mailback areas | 48,358 (1.5) | 5,728 (0.9) | 54,086 (1.4) |