

## RESULTS OF THE CENSUS 2000 BLOCK CANVASSING OPERATION

Joseph Burcham, U.S. Census Bureau  
4301 Suitland Road, Rm 1107/2, Suitland, MD 20746

**Key Words:** Master Address File, Block Canvassing, Blue-Line, Census 2000, City-style addresses

### Purpose

To provide a profile of addresses that the Census 2000 Block Canvassing operation added to, deleted from, and corrected on the Master Address File (MAF). I report only my major findings in this paper.

### Background on Block Canvassing

The Block Canvassing operation was one of the largest MAF building operations that the Census Bureau conducted for Census 2000. It occurred in the winter/spring of 1999.

It occurred in city-style address areas. A city-style address is an address that is composed of a house number and street name. City-style areas are areas where the majority of addresses are city-style. Non-city-style areas are areas that contain a large number of non-city-style addresses.

Block Canvassing occurred in parts of every state and in parts of 2,119 counties out of a total 3,143 counties in the nation. The roughly 1,000 counties where Block Canvassing did not occur were entirely non-city-style counties.

In the operation, field listers visited all of the city-style areas across the nation to update the residential addresses on the MAF. They carried listings of addresses from the MAF with them and compared the addresses on the listings to those that they observed. The listers recorded any necessary corrections, additions, and deletions on the listing pages. A total of about 92 million residential addresses were sent to be verified in Block Canvassing. These 92 million addresses represent about 60% of addresses on the MAF.

### Updating the MAF before Block Canvassing

The MAF is the Census Bureau's file containing residential addresses used for conducting the census and various surveys.

Several different sources were used to update the MAF in city-style areas. When reviewing some of the results in this paper, it is necessary to understand the sources that happened before Block Canvassing. These sources are:

- *the 1990 Census Address Control File (ACF)* - file of addresses collected in the 1990 census
- *the November 1997 Delivery Sequence File (DSF)* - file of addresses from the U.S. Postal Service
- *the September 1998 DSF, and*
- *\*the 1998 Local Update of Census Addresses (LUCA 98)* - address updates made to census files by local governments

\* LUCA 98 occurred before Block Canvassing only in some areas of the nation. In other areas, Block Canvassing updated the MAF before LUCA 98. In still other areas, Block Canvassing and LUCA used the same version of the MAF to provide updates (in other words, they happened at the same time).

### Updating the MAF with Block Canvassing Results

Block Canvassing results were used to classify each address as one of the following:

- *Verified* ..... 81,115,466 addresses
- *Adds* ..... 6,389,271 addresses
- *Deletes* ..... 5,146,320 addresses
- *Corrected Addresses* ..... 2,295,168 addresses
- *Geographic Corrections* .... 2,948,414 addresses

---

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

The *verified* category included all addresses that were verified as residential. The *adds* were residential addresses that the lister located but were missing from the listing pages. *Deletes* were such things as duplicates, addresses that did not exist in the lister’s search area, non-residential addresses, and uninhabitable addresses. *Corrected addresses* were verified as existing but required a correction to house number, street name, or a different address field. *Geographic Corrections* were verified as existing but required addresses with corrected geography codes.

There was an additional category that addresses were assigned to called “Add and Verify.” An “Add and Verify” address resulted from two different addresses that were recognized as referring to the same, where one address was originally an add and the other address was originally a verified address. For some of these addresses, the add and verify were originally in the same block. For others, the add and verify were in different blocks. The ones that started in different blocks could be thought of as geographic corrections. Considering this fact the true number of geographic corrections is between 2,948,414 and 3,055,816. However, we do not include the “Add and Verify” addresses in the remaining results in this paper.

When Block Canvassing listers took no action on an address, the Block Canvassing action code for that address was supposed to remain blank on the MAF. However, we received information that sometimes these blank values were converted to “V” (verified as existing unit) on the MAF. We do not know the magnitude of this occurrence. We do not have the ability to distinguish between addresses that were verified from addresses that received no action from Block Canvassing.

Reduction of Non-residential addresses and addresses with unknown block codes on the MAF

The Census Bureau selected residential addresses that were coded to Census 2000 collection blocks to be verified in Block Canvassing. Therefore, addresses with unknown block codes and non-residential addresses prior to Block Canvassing did not appear on the Block Canvassing listing pages. The Block Canvassing listers added any existing residential units that were missing from their listing pages. When the Census Bureau placed the Block Canvassing adds onto the MAF, we could recognize which of the addresses were already on the MAF but previously had an unknown block code or were coded as non-residential.

**What is the profile of Block Canvassing Adds?**

Total Adds - 6,389,271

This number represents a seven percent increase relative to the addresses initially sent to Block Canvassing to be verified.

The state level percent increase in Vermont is very high. Later in the paper I give a possible reason why this state level clustering occurred.

About 95 percent of the adds had city-style addresses. Because Block Canvassing occurred in the city-style areas, this result shows that the determination of areas to label as “city-style” was pretty accurate.

Block Code Agreement of Adds

The Census Bureau divides all of the land area in the nation into census blocks. In Block Canvassing (and other field operations), the listers were not only supposed to update the addresses on the MAF but also to code the addresses to the correct census block.

It was possible for different operations to provide different block codes for the same address on the MAF. When two or more operations provided block codes that disagreed, the Census Bureau used a scoring hierarchy to determine the official block code.

Table 1 shows the level of agreement between the Block Canvassing block code and the official block code.

**Table 1. Adds by Block Code Agreement**

<b>Block Canvassing block code was:</b>	<b># of Addresses</b>	<b>% of Total</b>
Same as official block	6,033,606	94.43
Different from official block	344,134	5.39
Not Provided	11,531	0.18
<b>Total Adds</b>	<b>6,389,271</b>	<b>100.00</b>

The block code provided by Block Canvassing was the same as the official block 94 percent of the time.

The “not provided” row represents addresses that did not receive a block code from Block Canvassing. It is possible for an address to be provided by several different sources but none of the sources provided a block code. These addresses stay on the MAF but have an unknown

block code.

### Original Source of Adds

Table 2 shows the magnitude of Block Canvassing adds in each of four original source categories.

**Table 2. Adds by Original Source**

Original Source	# of Addresses	% of Total
Pre-Block Canvassing	1,853,037	29.00
Block Canvassing	3,961,761	62.00
Block Canvassing and LUCA 98	568,915	8.90
Non-city-style Areas	5,558	0.09
<b>Total Adds</b>	<b>6,389,271</b>	<b>100.00</b>

The “Pre-Block Canvassing” category contains addresses that were first added to the MAF by an operation that occurred before Block Canvassing. This category contains about 29 percent of the adds.

At first look, this result is surprising because one would think that an address was *added* by Block Canvassing because it was not already on the MAF. The reason this happened is that these addresses either had an unknown block code or were coded non-residential on the MAF when the Block Canvassing universe was created. Addresses of these types were not sent to Block Canvassing to be verified. Block Canvassing added these addresses if they referred to good units and then the added addresses ended up merging to addresses already on the MAF. What Block Canvassing really did for these addresses was assign a block code to them and/or correct their residential status.

The “Block Canvassing” category shows the addresses first added by Block Canvassing only.

The “Block Canvassing and LUCA 98” category shows addresses first added by Block Canvassing and LUCA at the same time, so both sources get credit as the original source.

Earlier in the paper I state that Block Canvassing was only conducted in the city-style address areas. However, some addresses added by Block Canvassing had an original source that was only valid for the non-city-style areas and ended up being coded to the non-city-style areas. The “Non-city style area” category shows the

magnitude of these addresses. Basically, we think this happened because Block Canvassing listers went outside of their boundaries into non-city-style areas to add units when they shouldn’t have.

### Size of Basic Street Address for Adds

Table 3 shows the magnitude of adds in each of the different basic street address (BSA) size categories. A BSA is a house number / street name address that may include one unit or multiple units.

**Table 3. Adds by Size of BSA**

Size of BSA	# of Addresses	% of Total
blank	19,777	0.31
Single unit	4,106,666	64.27
2-4 units	1,220,453	19.10
5-9 units	447,102	7.00
10-19 units	450,673	7.05
20-49 units	107,321	1.68
50+ units	37,279	0.58
<b>Total Adds</b>	<b>6,389,271</b>	<b>100.00</b>

The largest category, by far, is single units (64% of the total adds).

The size of BSA variable that was available to us reflected the number of units at each BSA that were input to the census. Not all addresses input to the census were determined to be good addresses during the census. The distribution shown in table 3 could be a little different from the distribution as of the final census inventory.

### “In Census” Status of Adds

Not all of the Block Canvassing adds were determined to be good addresses in the census. About 78 percent of them were. The other 22 percent of adds were either:

- erroneously added by Block Canvassing, or
- existing at the time of Block Canvassing but not existing as of April 1, 2000 (census day)

## What is the profile of Block Canvassing Deletes?

Total Deletes - 5,146,320

This number represents 5.6 percent of addresses initially sent to Block Canvassing to be verified.

Both the state level percent increase of adds and percentage of the universe deleted in Vermont is a lot higher than the other states (22 percent and 17 percent, respectively). A possible reason for this result is the fact that a lot of addresses in Vermont were converted to E-911 addresses. The Block Canvassing listers were not allowed to change house numbers on their listing pages. So, when the house number of an address changed because of E-911 conversions, the listers were supposed to add an address with the new house number and delete the address with the old house number. So, a large amount of address conversions could generate a large amount of adds and deletes in this way.

### Original Source of Deletes

Table 4 shows the magnitude of Block Canvassing deletes by original source.

**Table 4. Deletes by Original Source**

Original Source	# of Addresses	% of Total
1990 ACF	3,624,027	70.42
November 97 DSF (or earlier)	989,848	19.23
September 98 DSF	77,925	1.51
LUCA 98	382,550	7.43
LUCA 98 and September 98 DSF	5,853	0.11
Other	66,117	1.28
<b>Total Addresses</b>	<b>5,146,320</b>	<b>100.00</b>

This table tells us that the majority of the Block Canvassing deletes (70 percent) originally were provided by the 1990 ACF. The two DSFs accounted for an additional 21 percent.

The “other” category on the graph contains the addresses in the non-city-style areas.

Table 5 shows the magnitude of addresses in the Block Canvassing universe that were deleted, broken down by

original source.

**Table 5. Magnitude of the Block Canvassing Universe Deleted, by Original Source (# in universe in parenthesis)**

Original Source	# of Deletes (# in universe)	% of universe deleted
1990 ACF	3,624,027 (75,183,729)	4.82
November 97 DSF (or earlier)	989,848 (14,579,494)	6.79
September 98 DSF	77,925 (785,640)	9.92
LUCA 98	382,550 (531,830)	71.93
LUCA 98 and September 98 DSF	5,853 (246,474)	2.37
Other	66,117 (285,603)	23.15
<b>Total Addresses</b>	<b>5,146,320</b> <b>(91,612,770)</b>	<b>5.62</b>

The highest percentage, by far, is the one for LUCA 98. A high percentage of addresses originally provided by LUCA 98 were deleted in Block Canvassing. These addresses are highly clustered in Cook County, Illinois. If we dropped this county from the analysis, the LUCA percentage would be much lower.

### Size of Basic Street Address for Deletes

Table 6 shows the magnitude of deletes in each of the different BSA size categories.

**Table 6. Deletes by Size of BSA**

Size of BSA	# of Addresses	% of Total
blank	17,691	0.34
Single unit	2,661,950	51.73
2-4 units	1,206,604	23.45
5-9 units	581,737	11.30
10-19 units	522,502	10.15
20-49 units	115,045	2.24
50+ units	40,791	0.79
<b>Total Addresses</b>	<b>5,146,320</b>	<b>100.00</b>

The table tells us that about 48 percent of all deletes occurred in multi-units. Looking back at the table of adds by size of BSA, only about 35 percent of adds occurred in multi-units. We believe there are so many more multi-unit deletes than multi-unit adds because of a duplication problem in multi-units.

I will give an example of this duplication problem. Suppose that for a given multi-unit address, the 1990 census gave us a record for unit 1 and a record for unit 2. The postal service gave us a record for the entire basic street address, with no unit designation. In this situation, the Block Canvassing lister would probably have deleted the record with blank unit designation and verified the other two. If this occurred a lot, we could see a net increase in deletes. So, we believe that Block Canvassing played a big part in repairing the duplication problem in multi-units.

“In Census” Status of Deletes

When Block Canvassing deleted an address, we required a second confirmation that the address should be deleted before we dropped it from the census. If there was no confirmation, the Census Bureau put these addresses into the census process. After the census was completed, we found that about 24 percent of the Block Canvassing deletes actually turned out to be residential addresses in Census 2000. These addresses were either:

- erroneously deleted by Block Canvassing, or
- on the MAF all along, not built in time for Block Canvassing (causing Block Canvassing to delete them), then they were built in time for the census, (which verified their existence).

**What is the profile of Block Canvassing Addresses Corrected?**

Total Addresses Corrected - 2,295,168

This number represents about 2.5 percent of addresses sent to Block Canvassing to be verified.

**What is the profile of Block Canvassing Geographic Corrections?**

Total Geographic Corrections - 2,948,414

This number does not include the “Add and Verify” addresses, which include geographic corrections. See section on “Updating the MAF with Block Canvassing Results” for a more thorough explanation of the “Add and Verify” addresses.

The 2,948,414 geographic corrections represents about 3.3 percent of addresses sent to Block Canvassing to be verified.

**Conclusions**

Block Canvassing was a large operation that provided a large number of updates to the MAF.

Block Canvassing not only played a large role in adding existing addresses to and deleting non-existing addresses from the MAF, but also in coding of addresses to the correct census block.

The operation played a part in deleting duplicate addresses in multi-units.

There were some errors in the updating. These errors included things such as listers crossing their boundaries to add units and listers erroneously deleting valid addresses.

Finally, there was some clustering of certain characteristics in the updating. This clustering was evident in results such as the high number of adds and deletes in Vermont.

**Additional Analysis**

In this paper, I presented the first stages in my data analysis of Block Canvassing. There is much more analysis that I have been working on or will be working on in the future.

Some additional analysis includes:

- more research into the reasons for state-level clustering
- research on clustering by number of addresses in a block (for example: Did most of the Block Canvassing adds occur in a few large blocks or were the adds scattered across all types of blocks?)
- calculations by various other characteristics.

### **Future Research**

All of the results presented give merely a snapshot of the updates taken by Block Canvassing on the MAF. In the future, our staff will be conducting extensive research to compare Block Canvassing results to the results of other MAF building operations and comparing the results of Block Canvassing to the results of the census.

In keeping the MAF as up-to-date as possible, the Census Bureau has been exploring the targeting of areas with certain characteristics as priority areas for updating the MAF. The clustering results in this presentation could be used as a first step in showing us how to conduct the MAF update targeting.

### **References**

Burcham, Joseph (2000), "Assessing the Quality of the Census 2000 Master Address File Using Earlier Evaluation Data," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Burcham, Joseph (1999), "The 1998 Master Address File Quality Improvement Program," internal memorandum from Diane F. Barrett through Ruth Ann Killion for Robert Marx, Bureau of the Census.

Bureau of the Census (1999), "Block Canvassing Lister's Instructions," guide for the field representatives.

Bureau of the Census (1999), "Program Master Plan: Census 2000 Block Canvassing Operation," internal memorandum.

Bureau of the Census (2000), "Program Master Plan: Census 2000 Master Address File," internal memorandum.

Killion, Ruth Ann (2000), "Study Plan for the Evaluation of the Block Canvassing Operation - F.5," internal memorandum for distribution list, Bureau of the Census.

Miskura, Susan M. (1999), "Validating Block Canvassing Deletes in LUCA '98 Field Verification," internal

memorandum through Preston Jay Waite for John H. Thompson, Bureau of the Census.

Treat, James. (1999), "Specification of the Decennial Master Address File Deliverability Criteria for Census 2000," internal memorandum from Howard Hogan to Susan Miskura and Robert Marx, Bureau of the Census.

Vitrano, Frank et. al. (2000), "Quality of Census 2000 Processes (Prototype)," internal memorandum from David C. Whitford for Howard Hogan, Bureau of the Census.