

# An Investigation into Expanding the CAUS Universe<sup>1</sup>

Bryan Schar, James Lawrence, Star Ying, Jim Hartman

U.S. Census Bureau  
Washington, DC 20233

## Abstract

The Community Address Updating System (CAUS) conducts targeted block listings to improve the coverage of the American Community Survey's sample frame. Blocks with certain address characteristics have traditionally been excluded from listing by the CAUS program because it was believed that data provided by other sources kept them current. However, results from the nationwide 2010 Census Address Canvassing operation suggested that some of these excluded blocks may benefit from being listed by the CAUS program. Therefore, a study was conducted to determine if the universe of CAUS blocks should be expanded. This paper will discuss the methods being used to determine if the CAUS universe should be expanded, some preliminary results of the investigation, and plans for future research.

## 1. Introduction

### 1.1. Master Address File – Description and Updating

The Master Address File (MAF) is a U.S. Census Bureau file that contains an up to date inventory of all known living quarters in the United States. It is the sole source of housing unit records for the American Community Survey's (ACS) sampling frame, and the main source of group quarters (GQ) information for the ACS's GQ sampling frame (Bates, 2011). Between Decennial Censuses, the U.S. Postal Service's (USPS) Delivery Sequence File (DSF) is the primary source of city-style<sup>2</sup> address updates for the MAF. The Demographic Area Address Listing operation, of which the Community Address Updating System (CAUS) targeted census block listing operation is a part, is the primary source of non-city-style<sup>3</sup> address updates (GEO, 2011b).

### 1.2. Delivery Sequence File

The DSF is the USPS's list of delivery points in the country, with a delivery point being a single mailbox or other place at which mail is delivered. It is assumed that city-style delivery points on the DSF represent the location of the associated living quarters. The DSF also contains information about non-city-style delivery points; however, this information is not used by the Census Bureau to update the MAF for the following reasons:

- The non-standard format of certain types of non-city-style addresses on the MAF, such as physical description only addresses, makes it difficult to match them to non-city-style delivery points on the DSF. Thus, updating the MAF with non-city-style delivery points from the DSF increases the risk of adding duplicate living quarters to the MAF.
- There is a potential that non-city-style delivery points on the DSF may be duplicates of city-style delivery points in areas that have undergone a conversion from non-city-style to city-style.
- It cannot be assumed that non-city-style delivery points on the DSF represent the location of living quarters. For instance, a PO Box does not represent the location of the living quarters of the person renting the P.O. Box.
- The information provided by the DSF for many non-city-style delivery points cannot be used to geocode (i.e., assign it to a census block) the address represented by it.

Due to these concerns, non-city-style delivery points on the DSF are not used to update the MAF. Instead, CAUS targeted block listing operations are a major source of address updates to the MAF in areas with a high proportion of non-city-style address (GEO, 2011b).

---

<sup>1</sup> This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

<sup>2</sup> Addresses that contain at least a house number and street name.

<sup>3</sup> Examples of non-city-style addresses are PO Boxes, rural routes, and description only addresses.

### 1.3. CAUS Listings

A census block, or block for short, is the smallest geographic entity for which the Census Bureau collects and tabulates data. When a block is listed, a field representative canvasses that block in order to identify all living quarters present within the block. The CAUS branch is responsible for selecting the blocks that are listed through the CAUS program.

When selecting blocks to list, the CAUS branch first considers the Address Characteristic Type (ACT) code assigned yearly by the U.S. Census Bureau's Geography Division (GEO). This code indicates both the style of MAF addresses in a block and the percent of addresses that can be matched to delivery points on the DSF (GEO, 2011a). Note that a block's ACT code may change from year to year based on changes in its MAF information. The CAUS branch classifies blocks into one of three groups based on their ACT code.

- Yes, or "Y", blocks are considered to be in the CAUS universe and eligible for listing.
- Maybe, or "M", blocks are considered to be in the margins of the CAUS universe.
- No, or "N", blocks are considered to be outside the CAUS universe.

Blocks in the "N" group contain either only city-style addresses, non-residential addresses, or no addresses. Blocks in the "M" group contain a mixture of city-style and non-city-style addresses where some of the addresses match to a DSF delivery point and at least 80 percent of the addresses are city-style. Blocks in the "Y" group contain either only non-city-style addresses, a mixture of city-style and non-city-style addresses where either none or all of the addresses match to a DSF delivery point, or a mixture of city-style and non-city-style addresses where some of the addresses match to a DSF delivery point and less than 80 percent of the addresses are city-style. Table 1 gives the ACT codes that belong to each group, and attachment A provides information about the types of addresses in each ACT code.

**Table 1: ACT Code Groups Used for CAUS Listings**

CAUS Group	ACT Codes
Y	M1, ME, MF, MG, M3, N1, N2, N3, P1, P2, P3, R1, R2, R3
M	MA, MB, MC, MD
N	B1, B2, B3, C1, C2, C3, Z0

### 1.4. Motivation for Research

Traditionally, blocks within the "M" and "N" groups were not considered for listing because it was believed that delivery point data provided by the USPS were sufficient to keep those blocks current. Results from the 2010 Census Address Canvassing operation (AdCan) conducted in 2009 provided listing data for blocks in these groups. These data allowed the CAUS branch to examine if blocks that were traditionally excluded from the CAUS universe need to be listed through the CAUS program to pick up new units. In addition, research conducted by the U.S. Census Bureau's Demographic Statistical Methods Division suggested there might be benefits to using an expanded definition of the CAUS universe based on the percent of MAF addresses that can be matched to a DSF delivery point (Kennel, 2010).

## 2. Research

### 2.1. Data Used

#### 2.1.1. 2010 Address Canvassing

One source of data for this research was the 2010 AdCan results. For this operation, census workers around the nation looked for every place where people live, stay, or could live or stay. They compared what they saw on the ground to what was shown on the MAF. Based on their observations, they verified, updated, or deleted addresses already on the MAF, and added addresses that were missing from it (ALOIT, 2012). Only those 2010 AdCan unit records that met all the criteria below were considered for this study:

- The unit returned from the 2010 AdCan with an action code that indicated it was a new unit added by the operation.
- The unit was in the final 2010 Census universe, that is, it was classified as a housing unit in the 2010 Census.

- The unit was included in the 2012 ACS main<sup>4</sup> housing unit frame universe.

Information from the 2012 main version of the ACS unit frame universe was used to filter, that is include or exclude, these records since the version of the MAF used in its creation was the first to include information about whether or not a unit was in the final 2010 Census universe.

### 2.1.2. CAUS Experimental Listings

Given the unique nature of the 2010 AdCan, the CAUS branch conducted experimental block listings of its own. One goal of these listings was to obtain a second source of data that could be used to determine if and where the CAUS universe should be expanded.

When a block is listed by CAUS, census workers look for every place where people live, stay, or could live or stay in that block. They compare what they see on the ground to what is shown on the MAF. Based on their observations, they verify, update, or delete addresses already on the MAF, and add addresses that are missing from it.

Blocks were selected for this experiment in March of 2010, which was approximately a year after the 2010 AdCan was conducted. The first step in selecting blocks for this study was to assign each block to a category. The category a block was assigned to was based on which of the three CAUS ACT groups (“Y”, “M”, “N”) it belonged to and which of five possible housing unit based groups it belonged to. The definitions of the five possible housing unit based groups are detail in Table 2 below. This resulted in each block being assigned to one of 15 categories as shown in Table 3.

**Table 2: Housing Unit Groups for CAUS Experimental Listings**

Housing Unit Group	Number of Pre-Listing Housing Units
1	3 or fewer
2	4-8
3	9-15
4	16-29
5	30 or more

Groupings based on housing units were used because initial attempts to model the expected number of new housing units, or adds, that were found during 2010 AdCan listings showed that the number of pre-existing housing units in a block was a major predictor of the number of adds that would be found when listing. Since the distribution of pre-listing housing units per block has a strong positive skew, including this variable in the models resulted in approximately 90 percent of the blocks (those with less than 30 housing units) being clustered together with very low predicted adds and the remaining 10 percent (those with 30 or more housing units) of blocks having high predicted adds. However, the CAUS branch wanted to be able to better distinguish between blocks with lower predicted adds rather than have them treated as essentially equivalent. Therefore, it was hoped that transforming pre-existing units into a categorical variable would help to provide more gradation to predicted adds.

The blocks within each of these groups with 100 or less pre-existing housing units were then sorted based on a selection score assigned to them. This score was derived from a generalized linear model that was developed using 2007 and 2008 block characteristic data to predict 2010 AdCan adds, and it indicated the expected number of new housing units, or adds, that would be obtained by listing the blocks. Only blocks with 100 or less pre-existing housing units were considered for this study due to the cost associated with listing blocks with a large number of pre-existing housing units. At the time of this study, approximately eight million of the 8.2 million tabulation blocks in the country had 100 or less pre-existing housing units.

A systematic sample of these sorted blocks was then selected from within each group. This was done to allow for a study of how well the selection score ranked blocks by adds, and how well it estimated the number of adds across a range of scores. The number of blocks selected from each group was driven by the various goals and constraints of the study. First, the CAUS branch estimated that it could afford to list around 7,500 blocks for this study. Based on previous experience, the CAUS branch only expected half as many blocks as it selects and sends out for listing to be completed. Therefore, it wanted to send

<sup>4</sup> Two versions of the ACS unit frame universe are created for each ACS survey year, the main and the supplemental.

out approximately 15,000 blocks in total for this study. Also, at the time that blocks were selected, the main interest was determining whether the “M” blocks should be added to the CAUS universe. The status of the “N” blocks was of secondary interest. Therefore, approximately four times as many “M” blocks were selected than “N”. In addition, the CAUS branch wanted to make sure that the listings resulted in a good number of new addresses being added to the MAF. Therefore, approximately half of the blocks it selected were “Y” blocks, and approximately three times as many blocks were selected from housing unit groups three, four, and five as compared to housing unit groups one and two.

This resulted in 15,241 blocks being selected and sent out for listing during April 2010 through June 2010. Of the blocks sent out, 8,498 were actually listed in the field. Tables 3 and 4 provide more details of this study’s block distribution.

**Table 3: Design of 2010 CAUS Experimental Listings**

Pre-Listing CAUS Group	Block Size Group	Blocks in Group		Blocks Sent for Listing		Blocks Listed	
		Count	Percent of Total	Count	Percent of Group	Count	Percent of Sent
Y	1	252,893	3.2	892	0.4	439	49.2
	2	158,925	2.0	892	0.6	440	49.3
	3	87,656	1.1	1,863	2.1	977	52.4
	4	62,347	0.8	1,862	3.0	951	51.1
	5	46,037	0.6	1,562	3.4	817	52.3
M	1	9,659	0.1	787	8.2	438	55.7
	2	62,596	0.8	792	1.3	437	55.2
	3	79,397	1.0	1,766	2.2	994	56.3
	4	88,030	1.1	1,768	2.0	978	55.3
	5	102,720	1.3	1,465	1.4	783	53.5
N	1	3,648,644	45.7	169	0.0	129	76.3
	2	871,375	10.9	169	0.0	128	75.7
	3	850,954	10.7	392	0.1	289	73.7
	4	934,728	11.7	392	0.0	299	76.3
	5	736,974	9.2	470	0.1	399	84.9
Total		7,992,935	100.0	15,241	0.2	8,498	55.8

Sources: U.S. Census Bureau, CAUS Targeting Database, February 2010; U.S. Census Bureau, CAUS Sample Control Output File, July 2010.

**Table 4: Distribution of 2010 CAUS Experimental Listings by ACT Code**

Pre-Listing CAUS Group	Pre-Listing ACT Code	Blocks in ACT		Blocks Sent for Listing		Blocks Listed	
		Count	Percent of Total	Count	Percent of Group	Count	Percent of Sent
Y	M1	106,398	1.3	1,629	1.5	812	49.9
	ME	50,441	0.6	700	1.4	355	50.7
	MF	22,437	0.3	436	1.9	237	54.4
	MG	219,141	2.7	3,028	1.4	1,563	51.6
	M3	133	0.0	4	3.0	3	75.0
	N1	131,929	1.7	935	0.7	466	49.8
	N2	53	0.0	1	1.9	0	0.0
	P1	21,224	0.3	115	0.5	56	48.7
	R1	56,092	0.7	223	0.4	132	59.2
	R2	8	0.0	0	0.0	0	
	R3	2	0.0	0	0.0	0	

Pre-Listing CAUS Group	Pre-Listing ACT Code	Blocks in ACT		Blocks Sent for Listing		Blocks Listed	
		Count	Percent of Total	Count	Percent of Group	Count	Percent of Sent
M	MA	87,325	1.1	1,587	1.8	851	53.6
	MB	109,513	1.4	2,096	1.9	1,180	56.3
	MC	74,685	0.9	1,376	1.8	769	55.9
	MD	70,879	0.9	1,519	2.1	830	54.6
N	B1	457	0.0	0	0.0	0	
	B2	156	0.0	0	0.0	0	
	B3	235,951	3.0	12	0.0	10	83.3
	C1	181,195	2.3	59	0.0	41	69.5
	C2	942,867	11.8	397	0.0	305	76.8
	C3	3,241,112	40.6	984	0.0	787	80.0
	Z0	2,440,937	30.5	140	0.0	101	72.1
Total		7,992,935	100.0	15,241	0.2	8,498	55.8

Sources: U.S. Census Bureau, Geography Reference File-Codes, July 2009; U.S. Census Bureau, CAUS Sample Control Output File, July 2010.

Only those unit records from these listings that met the criteria below were considered for the preliminary part of this study:

- The unit was accepted by GEO for inclusion on the MAF as a new address record.
- The unit was included in the 2011 ACS supplemental housing unit frame universe.

Information from the 2011 supplemental version of the ACS unit frame universe was used for these records since this is the first version of the ACS unit frame to include the addresses added by these listings.

Additional data used in this research comes from the MAF extracts and geography files used by ACS to create its unit frame universes each year. These data sources contain such information as a unit's DSF status, whether or not a unit is residential, and if the unit is considered to be within the ACS universe at the time of frame creation.

## 2.2. Decision Criteria

The primary function of CAUS listings is to serve as a complement to the updates provided by the DSF. This is why CAUS has traditionally only listed areas with a high proportion of pre-existing non-city-style addresses. However, if there was evidence that the DSF was not providing sufficient and/or timely updates to other areas, there could be a benefit from including them in the CAUS universe.

With this in mind, the CAUS branch decided to base its decision on whether or not to include new areas in the CAUS universe on a measure of how well DSF updating captures the adds from the 2010 AdCan and the 2010 CAUS experimental listings. In particular, if an address that was added to the MAF by one of these operations could not be matched by the Census Bureau to a delivery point on the DSF within a reasonable amount of time, this would be considered a deficiency in the ability of DSF updating to capture that add.

There are two aspects to this measure. First is the expected lack of DSF usability for capturing non-city-style adds. Though this deficiency is expected, since the Census Bureau does not use the DSF to add non-city-style addresses to the MAF, it still provides a measure of how much an area would benefit from being listed by the CAUS program. Second, is the unexpected lack of DSF usability for capturing city-style adds. Thus an area can be considered to have a poor DSF capture rate if most or all of its adds were non-city-style adds and/or were city-style adds that could not be matched to a delivery point on the DSF within a reasonable amount of time. Groups of blocks whose DSF match rates of adds are low enough, as measured by the percent of adds that can be matched to a delivery point on the DSF, may be considered for inclusion into the CAUS universe.

What constitutes a reasonable amount of time to wait for an added city-style address to match to a DSF delivery point is still an open question. It takes approximately a year, from beginning to end, for a block's information to be updated on the MAF through the CAUS program. This includes creating the frame from which to select the blocks, selecting blocks from this frame to list, having the selected blocks listed in the field, receiving the block updates from field, processing and editing this information, and finally applying the vetted updates to the MAF. During this same time period, the MAF is updated twice by the DSF (GEO, 2011b). If all the new city-style addresses obtained by listing a block through CAUS were also added by these two DSF updates, then there was no benefit, in terms of adds, gained by listing that block. However, what if it takes one year after the listing updates are applied for a city-style address to be captured using a DSF update? What about two years?

One important factor in determining this criterion is cost. The Census Bureau will continue to use DSF updates to keep a block's city-style address information on the MAF current regardless of whether or not that block is listed by the CAUS program. This means any costs associated with acquiring DSF updates are independent of the decision to list a block. Therefore, it will always be cheaper, in terms of dollars, to wait for city-style addresses to be added by DSF updates (when possible). However, other criteria, such as frame coverage, should be considered when making this assessment.

Since the traditional CAUS universe only includes areas with high concentrations of non-city-style addresses, and the DSF is not used to add these types of addresses to the MAF, the cost, in terms of frame coverage, of not listing these areas is clear. The properties of non-city-style addresses are different enough from those of city-style addresses that it would be difficult to justify imputing estimates for the former using information from the latter. However, if the CAUS universe is expanded to include areas that contain high concentrations of city-style addresses that, for reasons yet unknown, are not being added to the MAF by the DSF updates, then it is not as clear-cut that the coverage issues caused by this could not be accounted for through methods other than block listings. This suggests future research to investigate if city-style addresses being added to the MAF by the DSF are representative of those that are not.

## **2.3. Limitations**

### **2.3.1. Possible Overestimation of the AdCan Adds**

When census field representatives were assigned to canvass blocks during AdCan, they were provided with a list of the MAF addresses that had been geocoded to the assigned blocks. They were instructed to add addresses they found in the block that did not appear on that list and to delete any addresses from the list that they could not find in the block. This means that some of the adds in AdCan may have represented an address that existed on the MAF before AdCan but was either not geocoded to a block or was geocoded to the wrong block. Every effort was made during data processing to identify these false adds, but some of them could have been missed (ALOIT 2012, 115). This could result in an overestimate of the adds in the AdCan data, because some of the adds in AdCan were corrections to the geocoding of existing units.

In addition, some city-style addresses added by AdCan actually represented housing units that already existed as non-city-style addresses on the MAF that had been converted to city-style as part of enhancements to an area's emergency 911 system. Since census workers could have deleted the units' associated non-city-style address, these would form another type of add-delete pair that could have also slipped through the add-delete matching done during data processing. This could result in an overestimate of the city-style adds in the AdCan data, because some of the adds in AdCan were just conversions of existing units' addresses from non-city-style to city-style.

A complication introduced by the possible existence of false adds in the AdCan data is that they may behave differently, in terms of matching to DSF delivery points, than true adds. If this is the case, then inferences made about the DSF match rate of adds based on AdCan data may not be reflective of what would occur in regular CAUS listings.

### **2.3.2. ACT Code Representation in 2010 CAUS Experimental Listings Data**

As mentioned in Section 2.1., blocks were selected for the 2010 CAUS experimental listings based on their ACT group and housing unit group. This was done because it was felt that blocks that belonged to the same ACT and block size groups were similar in terms of adds. In addition, preference was given to selecting blocks in the "Y" and "M" ACT groups and blocks with more pre-existing residential housing units. A consequence of this is that some ACT codes were not represented in this study, while others only had a limited number of blocks selected. These ACT codes tended to occur rarely in the universe of blocks and/or be in the "N" ACT group.

### 2.3.3. Sparsity of Adds in Data

Reliably estimating the expected DSF match rate of adds for a group of blocks requires that enough adds were found in enough of its blocks during AdCan and/or the 2010 CAUS experimental listings. This can present a challenge since only 16.0 percent of the blocks listed during AdCan contained adds, and 31.5 percent of the blocks listed during the 2010 CAUS experimental listings contained adds. Also, those blocks with adds in AdCan only contained about 4.2 new units on average, while those blocks with adds in the 2010 CAUS experimental listings contained about 5.4 new units on average.

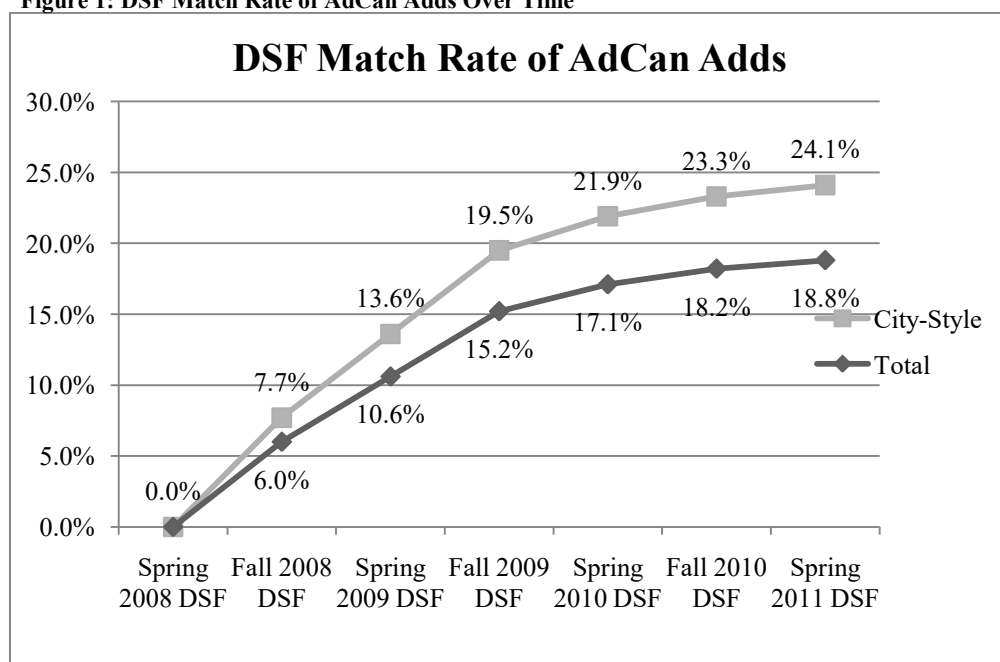
### 2.4. Results to Date

An initial examination of the change over time in the DSF match rate of adds from both AdCan and the 2010 CAUS experimental listings was conducted. Figures 1 and 2 give two measures of the percent of adds matched to the DSF for the AdCan and the 2010 CAUS experimental listings respectively. The square data points in each figure show the percent of city-style adds that have been matched to a delivery point across different vintages of the DSF. The diamond data points in each figure show the percent of total adds, both city-style and non-city-style, that have been matched to a delivery point across different vintages of the DSF. Note that, since non-city-style addresses are not matched to DSF delivery points (with only a handful of exceptions), the match rate for all addresses will always be less than or equal to the match rate for city-style addresses. The first DSF vintage shown on these figures, Spring 2008 for AdCan and Fall 2009 for the 2010 CAUS listings, is the one used as an input when creating the initial address lists for these operations.

As can be seen from Figures 1 and 2, the DSF match rate increases over time for both sets of data. However, the behavior of the percent adds matched to the DSF for AdCan adds suggests the rate of the increase may slow as time goes on. Further examination could be conducted to determine if there is a point where the marginal cost savings of waiting for an address to come in on the DSF is outweighed by the costs of excluding that unit from the sampling frame.

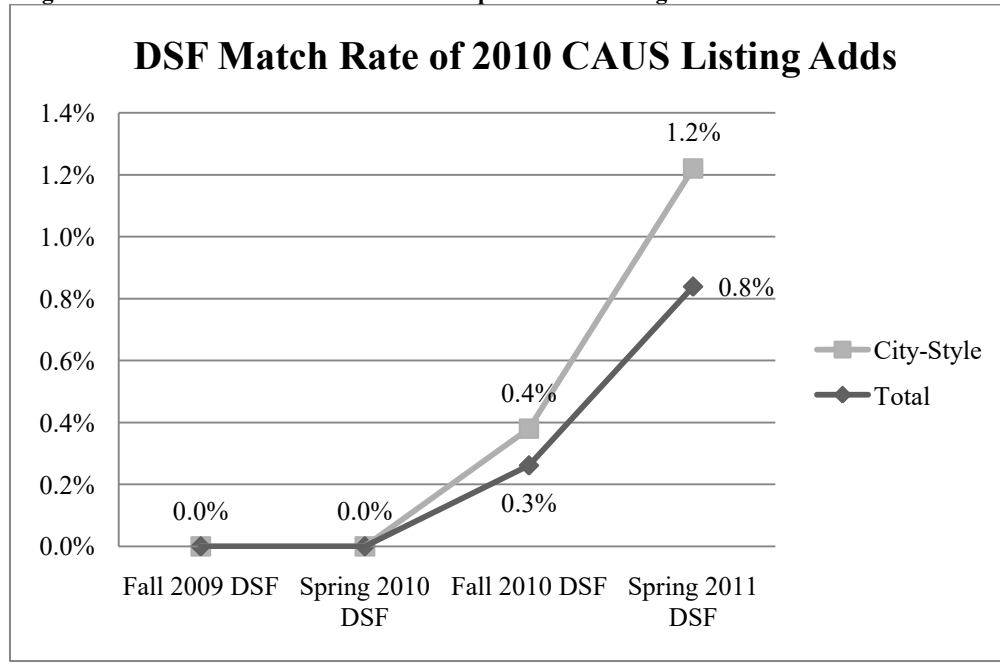
It is important to note that the match rates shown in Figures 1 and 2 are not cumulative, but represent how many adds were matched to a delivery point on each vintage of the DSF independent of whether or not they match to a delivery point on a previous vintage. There are cases of delivery points that stop appearing on DSF updates. This occurs when the USPS no longer considers an address to be a valid delivery point for mail. For example, when a previously occupied housing unit has been demolished. This means that the overall DSF match rate of adds could decrease from one vintage of the DSF to the next, though this has not yet been observed in this research.

Figure 1: DSF Match Rate of AdCan Adds Over Time



Source: U.S. Census Bureau, Master Address File, July 2011.

**Figure 2: DSF Match Rate of 2010 CAUS Experimental Listing Adds Over Time**



Source: U.S. Census Bureau, Master Address File, July 2011.

One thing that stands out from examining the DSF match rate of adds given in Figures 1 and 2, is how much higher the values are for the AdCan adds than the 2010 CAUS experimental listings adds. For example, the adds from AdCan had an overall DSF match rate of 15.2 percent from the Fall 2009 DSF, which represents three post listing DSF updates. The adds from the 2010 CAUS experimental listings have an overall DSF match rate of only 0.8 percent for an equivalent number of DSF updates. This could be due to the limitations of the AdCan data as discussed in Section 2.3. In addition, many of the blocks canvassed in AdCan had not been fully listed since the 2000 Census, and therefore, the adds from AdCan can represent up to ten years worth of changes. This means that, for some blocks, the DSF had up to twelve years to catch up to ground truth given by AdCan. By comparison, the 2010 CAUS experimental listing was conducted only a year after AdCan. This means the DSF had only one and a half years to catch up to the ground truth given by the CAUS experimental listing.

This initial analysis suggests that a conservative method of estimating the DSF match rate of adds for the purposes of excluding and including blocks in the CAUS universe is to use the most recent version of the DSF available each time the CAUS universe is created. The underlying assumption of this method is that as long as an add eventually comes in on a DSF it is worth the wait. This is the rule that was used for the preliminary investigation.

Since the traditional CAUS universe is defined in terms of the ACT code, this was the first type of block groupings that were examined. Tables 5 and 6 give the DSF match rate of adds grouped by their pre-AdCan and pre-2010 CAUS experimental listing ACT codes respectively. Pre-listing ACT code, rather than post-listing ACT code, of blocks were examined because that is the ACT code that is known at the time of selection. The values for the DSF match rate of adds in these tables are based on the most recent version of the DSF, the Spring 2011 vintage, available to the Census Bureau. Attachments B and C give the DSF match rate of adds by pre-listing ACT code for all post-operation vintages of the DSF.



**Table 5: DSF Match Rate of AdCan Adds by Pre-AdCan ACT Code**

Pre-AdCan CAUS Group	Pre-AdCan ACT	Blocks in ACT	Blocks with Adds	AdCan Adds			Adds on Spring 2011 DSF		
				City-Style	Non-City-Style	Total	City-Style Adds on DSF	Percent of City-Style Adds	Percent of Total Adds
Y	M1	115,534	65,248	252,423	125,559	377,982	16,100	6.4	4.3
	ME	50,110	24,843	74,718	35,805	110,523	10,857	14.5	9.8
	MF	22,281	14,554	52,697	26,676	79,373	6,681	12.7	8.4
	MG	213,379	113,539	447,057	247,330	694,387	51,524	11.5	7.4
	M3	159	83	1,104	14	1,118	33	3.0	3.0
	N1	141,587	52,647	87,218	114,810	202,028	12,503	14.3	6.2
	N2	61	45	121	178	299	35	28.9	11.7
	N3	1	0	0	0	0	0		
	P1	21,993	5,438	6,092	5,337	11,429	452	7.4	4.0
	R1	62,439	16,095	16,942	15,670	32,612	3,736	22.1	11.5
	R2	16	7	19	4	23	12	63.2	52.2
	R3	4	2	0	2	2	0		0.0
	Total	627,564	292,501	938,391	571,385	1,509,776	101,933	10.9	6.8
M	MA	106,272	67,150	359,935	83,086	443,021	99,673	27.7	22.5
	<b>MB</b>	<b>111,730</b>	<b>61,792</b>	<b>188,519</b>	<b>80,167</b>	<b>268,686</b>	<b>38,385</b>	<b>20.4</b>	<b>14.3</b>
	<b>MC</b>	<b>75,585</b>	<b>40,048</b>	<b>119,237</b>	<b>57,516</b>	<b>176,753</b>	<b>20,908</b>	<b>17.5</b>	<b>11.8</b>
	<b>MD</b>	<b>70,604</b>	<b>34,826</b>	<b>93,493</b>	<b>48,014</b>	<b>141,507</b>	<b>15,312</b>	<b>16.4</b>	<b>10.8</b>
	Total	364,191	203,816	761,184	268,783	1,029,967	174,278	22.9	16.9
N	B1	439	40	32	62	94	2	6.3	2.1
	B2	168	9	29	0	29	19	65.5	65.5
	B3	238,471	4,395	84,196	1,872	86,068	48,626	57.8	56.5
	<b>C1</b>	<b>189,406</b>	<b>44,415</b>	<b>98,449</b>	<b>17,282</b>	<b>115,731</b>	<b>6,295</b>	<b>6.4</b>	<b>5.4</b>
	C2	1,021,573	315,568	1,123,949	146,779	1,270,728	283,202	25.2	22.3
	C3	3,297,091	371,906	1,121,747	142,310	1,264,057	379,973	33.9	30.1
	Z0	2,466,679	80,816	215,683	60,453	276,136	51,277	23.8	18.6
	Total	7,213,827	817,149	2,644,085	368,758	3,012,843	769,394	29.1	25.5
Total		8,205,582	1,313,466	4,343,660	1,208,926	5,552,586	1,045,605	24.1	18.8

Sources: U.S. Census Bureau, Geography Reference File-Codes, July 2008; U.S. Census Bureau, Master Address File, July 2011.

Looking at Table 5, many of the blocks currently outside the CAUS universe have a DSF match rate of total adds that make them promising candidates for inclusion in the CAUS universe. In particular, blocks that had ACT codes of C1, MB, MC, and MD (which are bolded in Table 5) going into AdCan have an above average proportion of blocks with adds and below average DSF match rates for total adds. Other ACT codes had much smaller DSF match rates of adds than would be expected given the DSF coverage of existing units. For example, blocks with an ACT code of C3 contain only city-style addresses that can be matched to delivery points on the DSF. However only 33.9 percent of the city-style AdCan adds and 30.1 percent of the total adds were matched to a delivery point on the Spring 2011 DSF.

**Table 6: DSF Match Rate of 2010 CAUS Experimental Listings by Pre-Listing ACT Code**

Pre-Listing CAUS Group	Pre-Listing ACT	Blocks Listed	Blocks with Adds	Listing Adds			Adds on Spring 2011 DSF		
				City-Style	Non-City-Style	Total	Spring 2011 DSF	Spring 2011 DSF Percent CS Adds	Spring 2011 DSF Percent All Adds
Y	M1	812	325	740	608	1,348	2	0.3	0.1
	ME	355	148	253	216	469	8	3.2	1.7
	MF	237	101	158	175	333	2	1.3	0.6
	MG	1,563	682	1,431	1,773	3,204	10	0.7	0.3
	M3	3	0	0	0	0	0		
	N1	466	166	108	501	609	1	0.9	0.2
	P1	56	10	4	23	27	0	0.0	0.0
	R1	132	24	8	38	46	0	0.0	0.0
	Total	3,624	1,456	2,702	3,334	6,036	23	0.9	0.4
M	MA	851	293	4,102	346	4,448	44	1.1	1.0
	MB	1,180	324	677	284	961	8	1.2	0.8
	MC	769	233	578	273	851	6	1.0	0.7
	MD	830	210	304	227	531	8	2.6	1.5
	Total	3,630	1,060	5,661	1,130	6,791	66	1.2	1.0
N	B3	10	2	171	0	171	0	0.0	0.0
	C1	41	1	14	0	14	0	0.0	0.0
	C2	305	79	994	43	1,037	27	2.7	2.6
	C3	787	74	448	15	463	5	1.1	1.1
	Z0	101	7	5	22	27	1	20.0	3.7
	Total	1,244	163	1,632	80	1,712	33	2.0	1.9
Total		8,498	2,679	9,995	4,544	14,539	122	1.2	0.8

Sources: U.S. Census Bureau, Geography Reference File-Codes, July 2009; U.S. Census Bureau, Master Address File Transaction File, December 2010; U.S. Census Bureau, Master Address File, January 2011 and July 2011.

As mentioned previously, the overall DSF match rate of adds from the 2010 CAUS experimental listings is much lower than the overall DSF match rate of adds from AdCan. Table 6 shows that this is the case for each ACT code group as well as overall. This low level of DSF matching suggests there could be benefit from including all blocks in the CAUS universe. Examining the ACT groups that AdCan data suggested as candidates for inclusion into the CAUS universe, we see that only one of the 41 blocks in the C1 group that were listed had adds. This does not provide enough information to make a conclusion about the post AdCan DSF match rate of adds for this ACT group. All ACT code groups in the “M” CAUS group have a good number of blocks with adds and a low DSF match rates of adds. This, together with the AdCan data, provides evidence that MB, MC, MD, and possibly MA should be considered for inclusion in the CAUS universe.

## 2.5. Future Work

Preliminary investigations showed that blocks with certain ACT codes may benefit from being listed through the CAUS program. Future research will look into using additional block characteristics to refine what, if any, blocks are added to the CAUS universe. Some possible characteristics of interest are geographic location, number of pre-existing units, types of pre-existing units, percent of pre-existing units that are city-style, and percent of pre-existing units that can be matched to a delivery point on the DSF. When conducting this examination, the sparseness of blocks with adds in the data will be taken into consideration. If these investigations suggest groups of blocks to be added to the CAUS universe, additional experimental listings will be conducted as a confirmation. Similar investigations will be conducted for other geographic types such as block groups and tracts.

The CAUS branch will also look into whether or not the criteria currently used to filter adds in this study should be changed. Some possible alternatives to the current criteria used to filter the data for this research are:

- Using inclusion in or exclusion from the most recent version of the ACS housing unit frame universe to filter the data.
- Using the same criteria that are used to filter housing unit for inclusion in or exclusion from the list of existing MAF addresses provided to Census field representatives when they conduct listings.
- Using the same criteria that are used to filter housing units for inclusion in or exclusion from the addresses used when computing a block's ACT code.

The CAUS branch will also work with MAF stakeholders to determine how long it should wait for an address added by a listing to appear on the DSF when determining the DSF match rate of adds, and if non-listing techniques can be used to compensate for the issues caused by groups of blocks with poor DSF match rates for city-style address adds.

### 3. References

ALOIT (Address List Operations Implementation Team). (2012). "2010 Census Address Canvassing Operational Assessment Report." <http://2010.census.gov/2010census/pdf/2010%20Census%20Address%20Canvassing%20Operational%20Assessment.pdf>, January 17, 2012. Washington D.C.: U.S. Census Bureau.

Bates, Lawrence. (2010). "Editing the MAF Extracts and Creating the Unit Frame Universe for the American Community Survey (2011 Supplemental Phase)." Internally distributed software specification published as *DSSD 2012 American Community Survey Memorandum Series #ACS11-UC-4*, December 29, 2011. Washington D.C.: U.S. Census Bureau.

———. (2011). "Editing the MAF Extracts and Creating the Unit Frame Universe for the American Community Survey (2012 Main Phase)." Internally distributed software specification published as *DSSD 2012 American Community Survey Memorandum Series #ACS12-UC-1*, June 2, 2011. Washington D.C.: U.S. Census Bureau.

GEO (Geography Division). (2011a). "Address Characteristic Type Software Requirement Specification." Internally distributed software specification, version 1.5, last modified February 3, 2011. Washington D.C.: U.S. Census Bureau.

———. (2011b). "Delivery Sequence File Refresh Software Requirements Specification." Internally distributed software specification, version 1.8, last modified July 28, 2011. Washington D.C.: U.S. Census Bureau.

Kennel, Timothy and Joel Martin. (2010). "Summary Report for Evaluating MAF Content Quality After the 2010 Decennial Address Canvassing (Doc. #2010-4.0-G-16)." Internally distributed research report, version 1.0, November 15, 2010. Washington D.C.: U.S. Census Bureau.

# Attachment A

## Address Characteristic Type (ACT) Code Definitions

Two-character variable that describes the type of Master Address File (MAF) addresses in the block, and indicates how many of the block's MAF addresses can be matched to delivery points on the US Postal Service's Delivery Sequence File (DSF).

**Table A-1: ACT Code Definitions with CAUS Groups**

CAUS Group	ACT Code	Definition
Y	D1	Description only, MAF description only addresses cannot be matched to DSF addresses
	M1	City-style and noncity-style, no addresses matched to DSF
	ME	City-style and noncity-style, some addresses matched to DSF where the percent City-style is in [75, 80)
	MF	City-style and noncity-style, some addresses matched to DSF where the percent City-style is in [70, 75)
	MG	City-style and noncity-style, some addresses matched to DSF where the percent City-style is in (0, 70)
	M3	City-style and noncity-style, all addresses matched to DSF
	N1	Assorted noncity-style, no addresses matched to DSF
	N2	Assorted noncity-style, some addresses matched to DSF
	N3	Assorted noncity-style, all addresses matched to DSF
	P1	PO Box, no addresses matched to DSF
	P2	PO Box, some addresses matched to DSF
	P3	PO Box, all addresses matched to DSF
	R1	Rural Route, no addresses matched to DSF
	R2	Rural Route, some addresses matched to DSF
	R3	Rural Route, all addresses matched to DSF
M	MA	City-style and noncity-style, some addresses matched to DSF where the percent City-style is in [95, 100)
	MB	City-style and noncity-style, some addresses matched to DSF where the percent City-style is in [90, 95)
	MC	City-style and noncity-style, some addresses matched to DSF where the percent City-style is in [85, 90)
	MD	City-style and noncity-style, some addresses matched to DSF where the percent City-style is in [80, 85)
N	B1	Non-residential only, no addresses match to DSF
	B2	Non-residential only, some addresses matched to DSF
	B3	Non-residential only, all addresses matched to DSF
	C1	City-style, no addresses match to DSF
	C2	City-style, some addresses match to DSF
	C3	City-style, all addresses match to DSF
	Z0	No addresses

## Attachment B

This attachment provides detailed information about the DSF match rate of 2010 AdCan adds. Note that only AdCan adds which were both in the final 2010 Census universe and residential housing units within the 2012 ACS main frame were included in the counts below. The Spring 2008 DSF was used as an input into creating the AdCan frame, and the Spring 2010 DSF is the first DSF update to the MAF after AdCan adds were added to the MAF.

**Table B-1: DSF Match Rate History of AdCan Adds by Pre-AdCan ACT Code**

Pre-AdCan CAUS Group	Pre-AdCan ACT	Blocks with AdCan Adds	Total AdCan Adds	Spring 2008 DSF	Fall 2008 DSF		Spring 2009 DSF		Fall 2009 DSF		Spring 2010 DSF		Fall 2010 DSF		Spring 2011 DSF		Any DSF	
				AdCan Adds on DSF	AdCan Adds on DSF	Percent of AdCan Adds	AdCan Adds on DSF	Percent of AdCan Adds	AdCan Adds on DSF	Percent of AdCan Adds	AdCan Adds on DSF	Percent of AdCan Adds	AdCan Adds on DSF	Percent of AdCan Adds	AdCan Adds on DSF	Percent of AdCan Adds	AdCan Adds on DSF	Percent of AdCan Adds
Y	M1	65,248	377,982	0	4,025	1.1	8,036	2.1	11,102	2.9	12,401	3.3	14,216	3.8	16,101	4.3	16,193	4.3
	ME	24,843	110,523	0	3,152	2.9	5,871	5.3	8,224	7.4	9,375	8.5	10,270	9.3	10,857	9.8	10,900	9.9
	MF	14,554	79,373	0	1,913	2.4	3,481	4.4	4,955	6.2	5,601	7.1	6,143	7.7	6,681	8.4	6,711	8.5
	MG	113,539	694,387	0	14,555	2.1	26,179	3.8	37,871	5.5	42,646	6.1	47,881	6.9	51,527	7.4	51,775	7.5
	M3	83	1,118	0	4	0.4	33	3.0	33	3.0	33	3.0	33	3.0	33	3.0	33	3.0
	N1	52,647	202,028	0	3,285	1.6	6,520	3.2	9,288	4.6	10,128	5.0	11,537	5.7	12,503	6.2	12,535	6.2
	N2	45	299	0	6	2.0	19	6.4	25	8.4	27	9.0	34	11.4	35	11.7	35	11.7
	P1	5,438	11,429	0	108	0.9	192	1.7	292	2.6	352	3.1	426	3.7	452	4.0	455	4.0
	R1	16,095	32,612	0	1,035	3.2	2,006	6.2	2,718	8.3	3,028	9.3	3,380	10.4	3,736	11.5	3,741	11.5
	R2	7	23	0	11	47.8	11	47.8	12	52.2	12	52.2	12	52.2	12	52.2	12	52.2
	R3	2	2	0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
	Total	292,501	1,509,776	0	28,094	1.9	52,348	3.5	74,520	4.9	83,603	5.5	93,932	6.2	101,937	6.8	102,390	6.8
M	MA	67,150	443,021	0	33,323	7.5	59,597	13.5	83,373	18.8	91,715	20.7	96,638	21.8	99,673	22.5	100,079	22.6
	MB	61,792	268,686	0	12,698	4.7	22,210	8.3	30,685	11.4	34,048	12.7	36,679	13.7	38,385	14.3	38,505	14.3
	MC	40,048	176,753	0	6,365	3.6	11,530	6.5	16,221	9.2	18,144	10.3	19,790	11.2	20,908	11.8	20,984	11.9
	MD	34,826	141,507	0	4,631	3.3	8,322	5.9	11,850	8.4	13,290	9.4	14,499	10.2	15,312	10.8	15,361	10.9
	Total	203,816	1,029,967	0	57,017	5.5	101,659	9.9	142,129	13.8	157,197	15.3	167,606	16.3	174,278	16.9	174,929	17.0
N	B1	40	94	0	1	1.1	1	1.1	1	1.1	2	2.1	2	2.1	2	2.1	2	2.1
	B2	9	29	0	0	0.0	0	0.0	19	65.5	19	65.5	19	65.5	19	65.5	19	65.5
	B3	4,395	86,068	0	12,159	14.1	27,654	32.1	39,799	46.2	46,228	53.7	48,174	56.0	48,626	56.5	48,709	56.6
	C1	44,415	115,731	0	1,085	0.9	2,580	2.2	3,983	3.4	4,740	4.1	5,439	4.7	6,295	5.4	6,533	5.6
	C2	315,568	1,270,728	0	91,104	7.2	158,096	12.4	225,620	17.8	254,302	20.0	273,247	21.5	283,202	22.3	283,996	22.3
	C3	371,906	1,264,057	0	127,460	10.1	220,121	17.4	317,509	25.1	355,199	28.1	372,513	29.5	379,973	30.1	380,791	30.1
	Z0	80,816	276,136	0	17,547	6.4	28,172	10.2	41,515	15.0	47,917	17.4	49,950	18.1	51,278	18.6	51,431	18.6
	Total	817,149	3,012,843	0	249,356	8.3	436,624	14.5	628,446	20.9	708,407	23.5	749,344	24.9	769,395	25.5	771,481	25.6
Total		1,313,466	5,552,586	0	334,467	6.0	590,631	10.6	845,095	15.2	949,207	17.1	1,010,882	18.2	1,045,610	18.8	1,048,800	18.9

Sources: U.S. Census Bureau, Geography Reference File-Codes, July 2008; U.S. Census Bureau, Master Address File, July 2011.

## Attachment C

This attachment provides detailed information about the DSF match rate of adds from the 2010 CAUS experimental listings. Note that only adds which were residential housing units within the universe of the 2011 ACS supplemental frame were included in the counts below. The Fall 2009 DSF was used as an input into creating the frame for these listings, and the Spring 2011 DSF is the first DSF update to the MAF after the adds from these listings were added to the MAF.

**Table C-1: DSF Match Rate History of 2010 CAUS Experimental Listings by Pre-Listing ACT Code**

Pre-Listing CAUS Group	Pre-Listing ACT	Blocks with adds	Total Adds	Fall 2009 DSF		Spring 2010 DSF		Fall 2010 DSF		Spring 2011 DSF		Any DSF	
				Fall 2009 DSF	Fall 2009 DSF Percent All Adds	Spring 2010 DSF	Spring 2010 DSF Percent All Adds	Fall 2010 DSF	Fall 2010 DSF % All Adds	Spring 2011 DSF	Spring 2011 DSF Percent All Adds	Any DSF	Any DSF Percent All Adds
Y	M1	325	1,348	0	0.0	0	0.0	1	0.1	2	0.1	2	0.1
	ME	148	469	0	0.0	0	0.0	0	0.0	8	1.7	8	1.7
	MF	101	333	0	0.0	0	0.0	0	0.0	2	0.6	2	0.6
	MG	682	3,204	0	0.0	0	0.0	4	0.1	10	0.3	10	0.3
	N1	166	609	0	0.0	0	0.0	1	0.2	1	0.2	1	0.2
	P1	10	27	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
	R1	24	46	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
	Total	1,456	6,036	0	0.0	0	0.0	6	0.1	23	0.4	23	0.4
M	MA	293	4,448	0	0.0	0	0.0	4	0.1	44	1.0	44	1.0
	MB	324	961	0	0.0	0	0.0	2	0.2	8	0.8	8	0.8
	MC	233	851	0	0.0	0	0.0	3	0.4	6	0.7	6	0.7
	MD	210	531	0	0.0	0	0.0	4	0.8	8	1.5	8	1.5
	Total	1,060	6,791	0	0.0	0	0.0	13	0.2	66	1.0	66	1.0
N	B3	2	171	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
	C1	1	14	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
	C2	79	1,037	0	0.0	0	0.0	18	1.7	27	2.6	27	2.6
	C3	74	463	0	0.0	0	0.0	1	0.2	5	1.1	5	1.1
	Z0	7	27	0	0.0	0	0.0	0	0.0	1	3.7	1	3.7
	Total	163	1,712	0	0.0	0	0.0	19	1.1	33	1.9	33	1.9
Total		2,679	14,539	0	0.0	0	0.0	38	0.3	122	0.8	122	0.8

Sources: U.S. Census Bureau, Geography Reference File-Codes, July 2009; U.S. Census Bureau, Master Address File Transaction File, December 2010; U.S. Census Bureau, Master Address File, January 2011 and July 2011.