December 31, 2002

DSSD A.C.E. REVISION II MEMORANDUM SERIES # PP-51

MEMORANDUM FOR   Donna Kostanich
                                 Chair, A.C.E. Revision II Planning Group

From: Deborah Fenstermaker
                                 Chair, A.C.E. Revision II Duplicate Subgroup

Prepared by: Thomas Mule
                                 Decennial Statistical Studies Division

Subject: A.C.E. Revision II Results: Further Study of Person Duplication

The attached document provides the results of matching the full Enumeration (E) and Population (P) samples to the enumerations in the census.

A.C.E. Revision II Results: Further Study of Person Duplication

Thomas Mule
Decennial Statistical
Studies Division

Executive Summary

Evaluations of the March 2001 Accuracy and Coverage Evaluation (A.C.E.) coverage estimates indicated the A.C.E. failed to detect a large number of erroneous census enumerations. One type of these erroneous census enumerations is duplicate census enumerations; census enumerations included in the census two or more times. The A.C.E. was not specifically designed to detect duplicate census enumerations beyond the A.C.E. search area. For the Executive Steering Committee on Accuracy and Coverage Evaluation II (ESCAP II) analysis, Mule (2001) identified cases that in addition to being enumerated within the search area were also enumerated at a location outside the search area. Since you can only be a correct resident at one of the locations and based on our random sample, we expected that these Enumeration (E) sample cases duplicated to housing units would be coded correct one-half of the time. Feldpausch (2001) showed this expectation did not hold in her ESCAP II analysis.

For purposes of producing the A.C.E. Revision II estimates, the Further Study of Person Duplication (FSPD) used matching and modeling techniques to identify links between the full Enumeration (E) and Population (P) samples to census enumerations outside the search area. The census enumerations include enumerations 1) eligible for the E sample, 2) in group quarters and 3) in housing units reinstated in the census during the Housing Unit Duplication Operation (HUDO). We also matched to enumerations deleted from the census during the HUDO. We used statistical and exact matching methods to link records together. Using these results, we modeled the probability of records linked together being the same person.

What are the major methodological differences from the ESCAP II analysis?

The following are different:

- We assigned a probability that the linked records are the same person. For ESCAP II, we used a model weight approach based on a Poisson distribution in determining the estimates of duplication. For this analysis, we replaced the model weight with a method that computes a probability instead. For more information on this methodology, see Fay (2002a). This way we could use all of the duplicate results in the revised estimates instead of having to use a subset which Fay (2002b) had to do in analyzing erroneous enumerations for ESCAP II.

- We identified ways we could use statistical matching methods more. We used statistical matching during our first stage to be able to identify more possible duplicate households. We used the Expectation-Maximization algorithm to develop new sets of matching parameters for our second-stage matching. We used more information when modeling the links of the statistical process to determine if they were the same person.

What are the Major Results from the Further Study of Person Duplication?

Our analysis identified the following:

- Our study estimated 5,826,477 duplicates in the Census.

- Our study estimated there would have been 8,731,572 duplicates in the census if the HUDO had not been implemented. We include this estimate because the A.C.E. Revision II estimation uses both links to cases reinstated and deleted by the HUDO.

- For the E-sample Eligible universe, our improvements in computer matching identified 61.9 percent more duplicates within the cluster as compared to our ESCAP II results. We identified 1,173,344 million duplicates within the cluster as compared to 724,687 duplicates identified by the ESCAP II analysis.

- Overall, we identified 3,536,136 duplicates in the E-sample Eligible universe. This was 575,461 more duplicates than was found in the ESCAP II analysis. Most of these were found within the cluster or one ring of surrounding blocks.

- We identified 2,102,986 duplicates outside the surrounding blocks. This total is comparable to the 2,089,107 duplicates outside the surrounding blocks identified for the ESCAP II analysis. While we have roughly the same aggregate total, we believe we have done this by more accurately determining the duplication status of each case.

- While our study and the ESCAP II analysis estimated roughly the same aggregate total outside the surrounding blocks, the distribution of duplicates by geography has changed. Our study estimated more duplicates in the same county and fewer in a different state than the ESCAP II analysis. We believe this result is based on the improvements in the matching and modeling which we were able to implement in this analysis.

- Our estimate of housing unit to group quarter duplication (615,738) is similar to our ESCAP II estimate.

- The estimates to the Reinstated and Deleted universes are consistent with the results from the ESCAP II analysis.

What was our estimate of efficiency within the cluster?

We estimated efficiency by using the duplicates detected by the A.C.E. clerks as a benchmark. We estimated two efficiency measures. The first was the estimate using only the links to cases in the A.C.E. universe as was done by Mule (2001). Using this approach, the overall efficiency was 64.73 percent. Mule (2001) estimated an efficiency of 37.8 percent within the cluster for the ESCAP II analysis.

The second estimate used the cases in the A.C.E. universe and duplicates to the cases detected in the HUDO as was done by Fay (2002). Using this approach, we estimated an overall efficiency of 86.94 percent. Fay (2002) estimated an efficiency of 75.7 percent in his ESCAP II analysis.

Both methods showed that we were more efficient in identifying duplicates when there were two or more duplicates between the housing units. For this group, we were able to effectively utilize statistical matching techniques to identify these duplicates. When there was only one duplicate between the units, we had to rely on exact matching methods which limits the number of duplicates that we could detect.

Why did we decide not to use an efficiency adjustment?

To apply these adjustments based on duplicates within the cluster to duplicates detected outside the cluster requires the assumption that for the specified subgroups, the mechanism that is causing the duplicates within the cluster are similar for the duplicates outside the cluster. This assumption is debatable because duplicates within the cluster can be caused by misdelivery of forms or families living close together. As the geographic distance increases, the duplicates are more likely to be movers or children in joint-custody situations. Also there may be other variables like age or the type of response (Both Mail returns, One Mail/One Non-Mail, or Both Non-Mail) which can show differential efficiency. Including these variables could produce different adjustments than the ones used in this analysis. Based on concerns about the assumptions required, we decided not to adjust the estimates for efficiency.

What were the major results of the P sample matching across the country?

We identified a large number of P-sample nonmovers who were enumerated at another residence outside the one ring of surrounding blocks. These cases raise the question as to whether some of these people were truly residents of the cluster on April 1, 2000. Our results show that approximately half of these cases were nonmatches within the cluster. The A.C.E. Revision II developed a methodology to account for measurement error in the residence status of these cases in the revised estimates.

1. BACKGROUND

Evaluations of the March 2001 Accuracy and Coverage Evaluation (A.C.E.) coverage estimates indicated the A.C.E. failed to detect a large number of erroneous census enumerations. One type of these erroneous census enumerations is duplicate census enumerations; census enumerations included in the census two or more times. The A.C.E. was not specifically designed to detect duplicate census enumerations beyond the A.C.E. search area. For the Executive Steering Committee on Accuracy and Coverage Evaluation II (ESCAP II) analysis, Mule (2001) identified cases that in addition to being enumerated within the search area were also enumerated at a location outside the search area. Since you can only be a correct resident at one of the locations and based on our random sample, we expected that these Enumeration (E) sample cases duplicated to housing units would be coded correct one-half of the time. Feldpausch (2001) showed this expectation did not hold in her ESCAP II analysis.

For purposes of producing the A.C.E. Revision II estimates, this study used matching and modeling techniques to identify links between the full Enumeration (E) and Population (P) samples to enumerations outside the search area. The census enumerations include enumerations 1) eligible for the E sample, 2) in group quarters and 3) in housing units reinstated in the census during the Housing Unit Duplication Operation (HUDO). We also matched to enumerations deleted from the census during the HUDO. We used statistical and exact matching methods to link records together. Using these results, we modeled the probability of records linked together being the same person.

For more information on the Housing Unit Duplication Operation in Census 2000, See Nash (2000).

For more information on the analysis of person duplication for the ESCAP II decision not to adjust the census for uses other than redistricting in October, 2001, see Mule (2001).

For more information on using person duplicates to identify erroneous enumerations not accounted for in the March 2001 estimates, see Fay (2002b).

2. METHODS

This memorandum focuses on the results of the person duplication in Census 2000 and matching of the P-sample records to the enumerations collected for Census 2000 across the country.

For more information on the statistical or exact matching done to generate these estimates, see chapter 5 in Kostanich (2003).

1

2.1   How did we categorize the units in this analysis?

Our analysis classifies person records in the census into the following categories based on the following types of units:

Table 1: Categories of Census Units in this Analysis

Category	Description
E-sample Eligible[1]	Persons enumerated in housing units that were eligible to be selected for the Enumeration sample (E sample) for the Accuracy and Coverage Evaluation.
Reinstated	Persons enumerated in housing units identified to be potential duplicates by the Census Duplicate Housing Unit process. These housing units were ineligible for the E sample and the A.C.E. matching. The Duplicate Housing Unit process examined these cases and reinstated them into the census count.
Group Quarters	Persons enumerated in group quarters
Deleted	Persons enumerated in housing units identified to be potential duplicates by the Census Duplicate Housing Unit process. These housing units were ineligible for the E sample and the A.C.E. matching. The Duplicate Housing Unit process examined these cases and did not include these in the census count.

[1] Does not include Remote Alaska

2.2   How did we generate estimates?

For each link, we assigned sampling weights and duplication factors. For the estimates of matches for the nonmover residents in the P sample, we also included the residence probability of each case as determined for the March, 2001 estimates. We produced estimates by summing the products of these weights and factors for various categories of interest

For some analyses, we formed categories for:

- types of housing units (whether the housing unit was counted in the census or not)
- types of group quarters
- a geographic location of the duplicate

Appendix A documents these categories of census units.

For part of the analysis, we calculated percent duplication for two of the A.C.E. post-stratification variables: Race/Ethnicity domain and Age/Sex categories. The denominator for these estimates of the percent duplication was the number of data-defined persons in census housing units not including Remote Alaska. For part of these analysis we calculated the percent of P-sample nonmovers who matched in our search for these same two groups. The estimate of P-sample nonmovers in this percent uses the residence status and probability determined in A.C.E. production for the March, 2001 estimates. For these estimates for race/ethnicity domains

or age/sex categories, we used the characteristics of the E-sample or P-sample record. Appendix B documents the race/ethnicity domains and the denominator counts for each domain. Appendix C documents the denominator counts for each age/sex category.

For variance estimates, we used a simple jackknife methodology on the final A.C.E. cluster design. These variance estimates should be slight underestimates of the variances if they reflected the full A.C.E. cluster sampling plan.

2.2.1. Which weight did we use?

We used the Targeted Extended Search sampling weight.

2.2.2 What factors needed to be assigned to each link?

We assigned **two factors to each link**. The first factor was an unbiased probability of duplication for the link. This is also called a multiplicity factor. The second factor was a probability of duplication. It is the probability that the link represents true duplication.

2.2.3 How did we assign the first factor, the unbiased probability for the E sample?

The first factor was an unbiased probability of duplication or multiplicity factor for the link. A naive approach would think that each link should represent one duplicate. This would overestimate the amount of duplication when searching within the same universe (example: E-sample eligible to E-sample eligible). Here is one simple example why. Record A is a duplicate of record B. There is only one duplicate here. If both records are in the E sample, then we would make two links (A to B and B to A). Thus, we need to assign each link a probability of ½ to correctly estimate one duplicate. If we assigned a probability of 1 to each link, we would have incorrectly estimated two duplicates.

See Appendix D for more information on how we assigned the unbiased probability to each link.

2.2.4 How did we assign the first factor, the unbiased probability for the P sample?

All matches from the P-sample cases to the census enumerations received an unbiased probability or multiplicity factor equal to 1. This is different than the E sample factor because we only matched from the P-sample cases to the census enumerations.

2.2.5 How did we assign the second factor, the probability of duplication?

The second factor was the probability of duplication of the link. This is the probability of the two records linked together being the same person. We estimated this probability for the census duplicates and the P-sample matches from our study.

For more information on the modeling methodologies to determine the probability of duplication, please see chapter 5 in Kostanich (2003).

3. LIMITATIONS

- This type of analysis has been conducted nationally only once before. We have attempted to improve the estimates of duplication so they can be used to produce the A.C.E. Revision II estimates.

- We only conducted automated matching due mostly to time constraints; there was no clerical matching or field work to resolve unknown matches. Likewise, a conservative automated matching algorithm was used to ensure that we can be confident in our identification of duplicates.

- All duplicates identified by A.C.E. were clerically identified. Clerks were able to use more characteristics and look at the scanned census forms to determine duplicates. Because of our approach, our estimate of E-sample to E-sample duplication within the cluster compared to the A.C.E. estimate will be a conservative underestimate of the duplication within this universe.

4. RESULTS

4.1 What are the estimates of duplication?

Table 2 shows our estimates of duplication. We estimate duplicates by geographic distance and the type of census enumerations. We include two estimates of the total. The first shows the estimate of the duplicates in Census 2000. The second shows what the estimates of duplication would have been if the HUDO had not been implemented.

Table 2: Overall Estimates of Person Duplication

Geography		Type of Record in Census			Total (Records in Census)	Duplicates to Records Deleted During HUDO	Total (Records in Census + Records Deleted During HUDO)
		E-sample Eligible	GQ	Reinstate			
Within Cluster		1,173,344 (46,173)	76,381 (15,736)	1,058,548 (48,295)	2,308,273 (74,924)	1,967,199 (94,454)	4,275,472 (129,245)
Surrounding Block		259,805 (21,718)	25,373 (9,701)	24,751 (6,971)	309,929 (24,734)	678,355 (57,469)	988,284 (65,896)
Outside Surrounding Block	Same County	1,011,920 (24,292)	231,774 (39,795)	482,015 (27,797)	1,725,709 (55,097)	208,246 (20,789)	1,933,956 (59,590)
	Different County, Same State	563,270 (18,873)	190,417 (9,488)	88,331 (12,567)	842,018 (25,154)	35,111 (7,262)	877,129 (26,615)
	Different State	527,796 (23,744)	91,793 (7,093)	20,959 (17,316)	640,548 (31,433)	16,184 (4,902)	656,732 (33,930)
Total		3,536,136 (68,045)	615,738 (46,003)	1,674,604 (60,317)	5,826,477 (110,721)	2,905,096 (116,541)	8,731,572 (177,071)

Source: FSPD. Standard errors in parentheses.

Some highlights of these results:

• Our study estimated 5,826,477 duplicates in the Census.

• Our study estimated there would have been 8,731,572 duplicates in the census if the HUDO had not been implemented. We included this estimate because the A.C.E. Revision II estimation used links to cases reinstates and deleted by the HUDO.

• For the E-sample Eligible universe, our improvements in computer matching identified 61.9 percent more duplicates within the cluster as compared to our ESCAP II results. We identified 1,173,344 million duplicates within the cluster as compared to 724,687 duplicates identified by the ESCAP II analysis.

• Overall, we identified 3,536,136 duplicates in the E-sample Eligible universe. This was 575,461 more duplicates than was found in the ESCAP II analysis. Most of these were found within the cluster or one ring of surrounding blocks.

• We identified 2,102,986 duplicates outside the surrounding blocks. This total is comparable to the 2,089,107 duplicates outside the surrounding blocks identified for the ESCAP II analysis. While we have roughly the same aggregate total, we believe we have done this by more accurately determining the duplication status of each case.

5

- While our study and the ESCAP II analysis estimated roughly the same aggregate total outside the surrounding blocks, the distribution of duplicates by geography has changed. Our study estimated more duplicates in the same county and fewer in a different state than the ESCAP II analysis. We believe this result is based on the improvements in the matching and modeling which we were able to implement in this analysis.

- Our estimate of housing unit to group quarter duplication (615,738) is similar to our ESCAP II estimate.

- The estimates to the Reinstated and Deleted universes are consistent with the results from the ESCAP II analysis. We were expecting this. The reinstated and deleted cases were identified during the computer matching of the Census Housing Unit Duplication Operation. Since computer matching of person records was used in that operation, we were expecting our computer matching process this time to identify the same duplicates as in the ESCAP II analysis.

4.2   What is the efficiency of finding duplicates?

To assess the efficiency of finding duplicates, our benchmark is the duplicates detected by A.C.E. during the clerical matching. In the ESCAP II analysis, we used the entire set of duplicates of the E-sample Eligible cases within the cluster to assess efficiency. In hindsight, we realized that only a subset of duplicates is appropriate to do this assessment.

Table 3 shows the results of the within cluster matching of the A.C.E. clerks versus the FSPD computer matching for the E-sample Eligible universe. This table shows two areas which need to be accounted for when comparing the results of the two processes.

The first area is whether the duplication was within the same housing unit. The table shows that FSPD did not identify many of the duplicates within this group. Since FSPD was asked to provide information on duplication outside the A.C.E. search area, we did not focus on this type of duplication. In order to capture this more fully, we would have approached the matching and modeling differently to capture these types of duplicates.

The second area is cases with Insufficient Information for matching. For these E-sample records, the name is blank or incomplete or the name is complete but the person has only one characteristic. These cases were determined to be insufficient information for A.C.E. matching and were not part of the duplicate search by A.C.E. clerical staff. However, these cases were part of the FSPD matching. We were able to identify 11,000 duplicates for these cases. We were able to do this based on using the matching results for the other persons in the housing units.

To determine the within cluster efficiency of the FSPD duplication as compared to A.C.E., we excluded the duplicates from these two areas. In the next section, we will compare the 1,148,555 duplicates detected by FSPD to the 1,774,421 duplicates detected by A.C.E.

6

Table 3: Within Cluster Results for A.C.E. Clerical Search and FSPD

Type of Duplicates	Type of E-sample Case	A.C.E.	FSPD
Different Census Housing Unit	Insufficient Information	0 (0)	10,957 (2,055)
	All Others	1,774,421 (58,155)	1,148,555 (45,709)
Same Census Housing Unit	Insufficient Information	0 (0)	2,395 (884)
	All Others	142,836 (9,533)	11,437 (3,285)
Total		1,917,258 (59,236)	1,173,344 (46,286)

Standard errors in parentheses.

Fay (2002) discussed the issue of measuring the efficiency of duplication within the search area. His work examined two methods for measuring efficiency. The first method takes into account only the computer matching of the A.C.E. universe which was used in Mule (2001). Fay (2002) suggests an alternative measure which accounts for not only the duplication detected in the A.C.E. population but also to the housing units reinstated or deleted in the HUDO in Census 2000. This measure can serve as an upper bound on the efficiency of the computer matching. In these results, we estimated the efficiency using both measures.

Fay (2002) estimated the alternative measure at an overall level. In these results, we wanted to investigate the efficiency for various subgroups. We formed subgroups by identifying the following characteristics related to efficiency

- Household Size categories and
- Number of Links between the Units (Whole HH, Partial but not All, or Only 1).

We have formed three household size categories which classify the number of people in each unit.

- 1 Person to 1 Person: We identified a duplicate link where both housing units only have one person.

- 1 Person to 2+: We identified a duplicate link where one of the housing units has only 1 person. The other housing unit has two or more persons. An example of this kind of duplicate is a 23 year old who used to live with her parents (2+ household) and then moved out on her own (1 person).

- 2+ to 2+ : We identified a duplicate link where both housing units have two or more people.

For the duplication between units with two or more people, we wanted to summarize the number of links between the two units. In this analysis, we formed three categories:

- Whole household: The number of links between the two units is equal to the number of people in the E-sample housing unit.

- Partial but not All: We have two or more links between the units but we have not identified all of the persons in the E-sample housing unit as being duplicated.

- Only 1: We have only one link between these two housing units.

Table 4 shows these results by Household size and Number of links between the units. These results show that we are more efficient identifying whole household or partial duplication for households with two or more persons. We expected these results based on using statistical matching during the first stage and how we modeled the probability of duplication when we had two links between the units.

The efficiency is not as high for identifying duplicates for the following groups:

- between one person households
- a duplicate between one person households and two-plus households, and
- when only one person was duplicated between two-plus person households.

Again, we expected these results because of having to rely on exact matching methods to be able to identify duplicates in these three groups.

Table 4 shows that the efficiency measure increases for the following groups when the alternative efficiency measure using links to reinstates and deletes is used:

- whole household duplication between two-plus person housing units
- partial household duplication between two-plus person housing units and
- between one person households.

The reinstated and deleted cases are from the HUDO. A component of this operation was the use of person matching to identify housing units which were on our list multiple times. These types of duplicates have more impact on the efficiency for whole-household and partial household duplication situations. The duplicates to reinstates and deletes are not related to one person being duplicated in two different housing units.

8

Table 4: Efficiency Estimates Within Cluster

HH Size	Number of Links	ACE Within Cluster	FSPD Estimate	FSPD Efficient (%)	Denominator[1]	FSPD Estimate	FSPD Efficient (%)
			Including Links to Reinstates and Deletes				
			No			Yes	
1 Person to 1 Person	Only 1	204,604 (10,055)	16,756 (2,882)	8.19 (1.35)	393,295 (14,075)	205,447 (10,299)	52.24 (1.79)
1 Person to 2+	Only 1	139,038 (8,146)	36,271 (3,744)	26.09 (2.48)	143,009 (8,254)	40,243 (3,957)	28.14 (2.52)
2+ to 2+	Whole HH	952,280 (39,696)	747,682 (31,772)	78.51 (2.37)	3,362,979 (92,348)	3,158,382 (90,191)	93.92 (0.77)
2+ to 2+	Partial (2+)	329,631 (21,963)	318,557 (19,666)	96.64 (5.67)	741,709 (39,919)	730,636 (39,003)	98.51 (2.55)
2+ to 2+	Only 1	148,869 (8,579)	29,288 (3,356)	19.67 (2.13)	150,722 (8,620)	31,141 (3,494)	20.66 (2.18)
Total		1,774,421 (53,349)	1,148,555 (43,185)	64.73 (1.27)	4,791,715 (119,603)	4,165,848 (114,677)	86.94 (0.58)

[1] The denominator of the Fay alternative is the A.C.E. estimate plus the FSPD estimate of duplicates to reinstates and deletes. Standard errors in parentheses.

Table 4 shows the resulting efficiency estimates for estimating person duplication within the cluster. Now, this raises the question of efficiency outside the cluster where we have no benchmark to assess this. To apply these adjustments based on duplicates within the cluster to duplicates detected outside the cluster requires the assumption that for the specified subgroups the mechanism that is causing the duplicates within the cluster is similar for the duplicates outside the cluster. This assumption is debatable because duplicates within the cluster can be caused by misdelivery of forms or families living close together. As the geographic distance increases, the duplicates are more likely to be movers or children in joint-custody situations. Also there may be other variables like the age or the type of response (Both Mail returns, One Mail/One Non-Mail, or Both Non-Mail) which can show differential efficiency. Including these variables could produce different adjustments than the ones used in this analysis.

In order to examine this assumption, Appendix E shows the effect of using these efficiency measures to adjust the estimate of duplicates for the E-sample Eligible universe. Each figure shows the estimates of duplication for one of the subgroups identified in Table 4. We show the estimates as the geographic distance between the duplicates widens. In each figure, we present three estimates. The first is our estimate of duplication. This has no adjustment for efficiency. The second is the estimate using the efficiency adjustment without the links to reinstates and deletes. The third is the estimate using the efficiency adjustment with the links to reinstates and

deletes. The adjustment is the inverse of the efficiency rate. For example, the efficiency rate of 19.67 percent translates into a multiplicative adjustment of 5.08.

Figure E5 shows the effect for the only one duplicate between two-plus household situation. We see that using an efficiency measure of roughly 20 percent leads to large increases in the estimates of duplicates. Figure E5 shows that our estimate of roughly 350,000 in the same county outside the surrounding blocks would be increased to roughly 1.5 million. We see similar large increases in Figure E5 for different county and different state results as well. This type of duplication could be joint-custody children. Since we don't expect to find to find this type of duplication within a very small geographic area like a block cluster, it brings into question the appropriateness of using this efficiency measure within cluster to adjust duplicates detected outside.

Figures E1 shows the effects for duplication between one person households. Figure E2 shows the effects for duplication between one person and two-plus person households. Both of these had low efficiency measures in Table 4. We see that these two adjustment also increase the estimates of duplicates. Because of the high measure of efficiency for whole household and partial household duplicates, the increase in Figures E3 and E4 for these respective groups is not that large.

Based on these concerns, we decided not to use an efficiency adjustment in the A.C.E. Revision II estimates.

4.3   **What are the patterns of duplication by Race/Ethnicity domains**?

Tables F1 and F2 show the percent duplication for the Race/Ethnicity domains by geography. Table F1 shows the total results using all of the duplicates in the Census identified in our analysis. Table F2 shows the results not including the duplicate links to the reinstated units. Both tables show that the Non-Hispanic Black and Hispanic duplicates outside the cluster and surrounding blocks are concentrated in the same county.

Table F3 shows the duplication by the type of group quarters. The Non-Hispanic Black domain had higher percent of duplication than the Hispanic domain between 1) housing units and correctional facilities and 2) housing units and college dorms. The Non-Hispanic Asian domain had a large percent of duplication to group quarters in college dorms. However, the estimates for the Non-Hispanic Asian domain also have a large variance.

Table F4 shows the percent duplication to persons in housing units which were deleted from the Census by HUDO. We see similar results for the Non-Hispanic White and Other, Non-Hispanic Black, Hispanic and Non-Hispanic Asian domains. The results for the Hawaiian and Pacific Islander, American Indian on AIR and American Indian off AIR are more variable because of the smaller sample sizes.

4.4   What are the patterns of duplication by Age/Sex categories?

Tables F5 and F6 show the percent duplication for the Age/Sex categories by geography. Table F5 shows the total results using all of the duplicates in the Census identified in our analysis. Table F6 shows the results not including the duplicate links to the reinstated units. Both tables show the similar patterns of duplication outside of the surrounding blocks. Duplication of persons under 30 years old is concentrated more in the same county while duplication of persons 50 years and older are concentrated more in a different state.

Table F7 shows the percent duplication for the Age/Sex categories by the type of group quarters. The table shows the 18-29 female duplication was predominantly in college dorms while the 18-29 males were duplicated in college dorms, correctional facilities and military group quarters.

Table F8 shows the percent duplication to persons in housing units which were deleted from the Census by HUDO. The table shows similar results across the age/sex categories.

4.5   What are the patterns of duplication by tenure?

Tables F9 and F10 show the estimates of duplication by tenure for different geographic distances. Since each duplicate link has two records, we classified links as both owners, one owner/one renter or both renters. Table F9 shows the total results using all of the duplicates in the Census identified in our analysis. Table F10 shows the results not including the duplicate links to the reinstated units. Both tables show the similar patterns of duplication outside of the surrounding blocks. Outside the surrounding blocks in the same county, we identify roughly the same number of duplicates being both owners and one owner/one renter. For different states, we identify more duplicates where both were owner than one owner/one renter or both renters.

Table F11 shows the duplication for the tenure categories by the type of group quarters. Since these are duplicates from housing units to group quarters, we show the duplicates by the owner and renter status of the housing unit. Over ½ of the owner duplicates to group quarters are to college dormitories as compared to roughly one-third of the renter duplicates. Almost twenty percent of the renter duplicates are duplicated to correctional institutions as compared to just over ten percent of the owner duplicates.

Table F12 shows the percent duplication to persons in housing units which were deleted from the Census by HUDO. We see that the duplicates to this universe decrease as geographic distance gets larger. This is because of the close geographic search area of HUDO in Census 2000. We also see that most of these duplicates are both owners or both renters. Since these housing units were determined to be duplicates by HUDO, we would expect any of the people duplicated between the two units would report the same tenure status each time.

11

4.6   What are the results of matching the P sample across the country?

Our matching of the E sample across the country allowed the A.C.E. Revision II methodology to adjust the correct enumeration rate to account for erroneous enumerations not detected in the production for the March, 2001 estimates or in the evaluation follow-up. To be consistent, we matched the persons in the P sample across the country to provide information to allow the match rate to be adjusted for errors in determining the residence status. Since the revised estimation methodology only uses the links of nonmovers determined to be residents in A.C.E. production for the March, 2001 estimates, we only present the results for this group in this analysis.

Table 5 shows the results of matching the nonmover residents of the P-sample to the census across the country. These cases were determined to be residents of the housing unit in the cluster on April 1st and did not move out before the A.C.E. interview. For this analysis, we used the residence and mover status determination of the original A.C.E. data. We divided the P-sample nonmovers by whether they matched to a census enumeration within the search area in A.C.E. production for the March, 2001 estimates. Table 5 shows that we were able to locate many P-sample nonmatches with an enumeration at another unit outside the surrounding blocks.

Table 5: Overall Results of Matching the Nonmover Residents

	Type of Record							
	E-sample Eligible		GQ		Reinstate		Deleted	
	Match Status of P sample		Match Status of P sample		Match Status of P sample		Match Status of P sample	
	Nonmatch	Match	Nonmatch	Match	Nonmatch	Match	Nonmatch	Match
Geography								
Within Cluster	416,280 (17,506)	199,026,173 (2,078,493)	0 (0)	92,379 (22,905)	473,167 (57,598)	912,493 (45,194)	242,867 (33,394)	2,050,732 (117,371)
Surrounding Block	512,407 (40,315)	8,886,048 (547,289)	5,158 (2,874)	4,118 (1,668)	50,725 (13,974)	61,334 (14,600)	26,104 (7,477)	323,939 (30,050)
Outside Surrounding Blocks — Same County	2,059,658 (116,361)	1,194,385 (34,618)	39,927 (8,720)	127,393 (25,135)	12,843 (3,963)	195,517 (17,458)	56,759 (24,401)	96,294 (13,639)
Different County, Same State	403,823 (28,067)	651,502 (23,513)	29,868 (4,155)	86,527 (6,467)	3,791 (1,732)	39,092 (7,308)	7,676 (3,455)	10,575 (2,928)
Different State	268,031 (19,922)	843,350 (24,656)	15,480 (2,312)	102,439 (6,299)	3,851 (2,348)	3,272 (839)	2,871 (1,017)	10,071 (2,574)
Total	3,660,200 (132,526)	210,601,459 (2,192,069)	90,433 (10,535)	412,855 (35,536)	544,376 (59,711)	1,211,708 (51,211)	336,277 (43,085)	2,491,612 (124,742)

Source: FSPD. These estimates include the residence probability. For this table, a case was considered a match if the probability of being match was greater than zero. We used the residence probability and match probability from the March, 2001 estimates. Standard errors in parentheses.

4.7 What are the matching patterns of nonmovers by Race/Ethnicity domains?

Table G1 and G2 shows the results of matching the nonmovers residents across the country. Table G1 shows the match results linking to all of the enumerations in the Census. Table G2 shows the results not including the links to the reinstated units. From our analysis, the Non-Hispanic White or Some Other Race domain has a larger total percent of matches than the Non-Hispanic Black or Hispanic domains. However for the same county outside the surrounding blocks, we see that the Non-Hispanic Black and Hispanic domains have a higher percent than the Non-Hispanic White and Some Other Race domain. We see that the Hispanic domain has a larger percent than the Non-Hispanic Black and the Non-Hispanic White and Some Other Race domain for matches to a different state.

Table G3 shows the matching of matching to type of group quarters. Similar to the duplicates to group quarters, the Non-Hispanic Black domain had higher amounts of matching to this universe than the Hispanic domain. The Non-Hispanic Asian domain had a large percent of matches to

group quarters in college dorms. However, the estimates for the Non-Hispanic Asian domain have a large variance.

Table G4 shows the results of matching to persons in housing units which were deleted from the Census by the HUDO. We see similar results for the Non-Hispanic White and Other, Non-Hispanic Black, Hispanic and Non-Hispanic Asian domains. The results for the Hawaiian and Pacific Islander, American Indian on AIR and American Indian off AIR are more variable because of the smaller sample sizes.

4.8 What are the matching patterns of nonmovers by Age/Sex Categories?

Tables G5 and G6 show the results of matching for the Age/Sex categories by geography. Table G5 shows the total results using the links to the enumerations in the Census. Table G6 shows the results not including the links to the reinstated units. Both tables show the percent of matches for persons under 30 years old is concentrated more in the same county for those outside the surrounding block.

Table G7 shows the results of matching for the Age/Sex categories by the type of group quarters. The tables shows that the percent of matches for 18-29 males were distributed among correctional institutions, college dorms and military group quarters. The table shows that most of the 18-29 females are matches to college dorms.

Table G8 shows the results of matching to persons in housing units which were deleted from the Census by HUDO. The table shows similar results across the age/sex categories.

References

Fay, R., "Probabilistic Models for Detecting Census Person Duplication," Proceedings of the Survey Research Methods Section, American Statistical Association, 2002a.

Fay, R., "Evidence of Additional Erroneous Enumerations from the Person Duplication Study," Executive Steering Committee on Accuracy and Coverage Evaluation Policy II Report 9, U.S. Census Bureau, Washington, D.C., 2002b.

Feldpausch, R., "Census Person Duplication and Corresponding A.C.E. Enumeration Status," Executive Steering Committee on Accuracy and Coverage Evaluation Policy II Report 6, U.S. Census Bureau, Washington, D.C. 2001.

Haines, D., "Accuracy and Coverage Evaluation Survey: Final Post-Stratification Plan for Dual System Estimation," DSSD Census 2000 Procedures and Operations Memorandum Series Q-24, U.S. Census Bureau, Washington, D.C., 2000.

Kostanich, D., "Chapter 5: Further Study of Person Duplication in the Census," DSSD A.C.E. Revision II Memorandum Series PP-30, U.S. Census Bureau, Washington, D.C., 2003.

Mule, T., "Person Duplication in Census 2000," Executive Steering Committee on Accuracy and Coverage Evaluation Policy II Report 20, U.S. Census Bureau, Washington, D.C., 2001.

Nash, F., "Overview of the Duplicate Housing Unit Operations," Internal Census Bureau memorandum, Census 2000 Informational Memorandum Number 78, U.S. Census Bureau, Washington, D.C., 2000.

Appendix A: Analysis Categories

A.1 Geographic Categories of the Duplicate Links

Our analysis used the following categories of geography:

Cluster and Surrounding blocks

- Within the block cluster
- With the one ring of surrounding blocks outside the cluster

Outside the cluster and one ring of surrounding blocks

- Within the same county
- Within a different county in the same state
- In a different state

A.2 Categories of Housing Units

Table A1 shows the categories of housing units in this analysis

Table A1: Categories of Housing Units

Category of Housing Units	Type of Units in Category
Census Housing Unit	• E-sample Eligible Housing Units • Reinstated
Deleted Housing Units	• Housing units removed during the Duplicate Housing Unit Operation

A.3 Categories of Group Quarters

Table A2 shows the categories of group quarters in this analysis.

Table A2: Categories of Group Quarters

Category of Group Quarters	Type of Units in Category
Correctional Institution	• Federal detection centers • Federal prisons • State prisons • Local jails • Correctional halfway houses • Military prisons • Other prisons
Nursing Homes	• Nursing home
Juvenile Institution	• Neglected/abused juvenile institutions • Emotionally distributed kids institutions • Delinquent kids institutions • Other juvenile institutions
College Dorms	• College dorms
Military	• Miliary barracks
Other	• Drug/alcohol abuse treatment • Military hospital • Civilian hospital • Hospices • Mentally ill hospital • Mentally handicapped hospital • Institution for deaf • Institution for blind • Other physically handicap • Homeless shelter • Children's shelter • Domestic violence shelter • Soup kitchen • Mobile food van • TNSOLs • Drug/alcohol group home • Mentally ill group home • Physically handicapped group home • Other group home • Agricultural worker's dorm • Other worker dorm • Job corps dorm • Staff dorms: Military hospital/prison • Religious group quarter • Hostels, YM/WCAs, etc. • Protective oversight

Appendix B: Race/Ethnicity Domains

The race/origin domain assignment generally follows the guidelines listed below, but it is essential to see Haines (2000) for the complete set of rules used to classify people into one of the seven domains. The race/origin domain assignment is hierarchical.

Domain 1 (American Indian or Alaska Native on reservations) includes:

- All people on a reservation with American Indian or Alaska Native either as their single race or as one of multiple races, regardless of their Hispanic origin.

Domain 2 (American Indian or Alaska Native off reservations) includes:

- All people in Indian Country[1] but not on a reservation with American Indian or Alaska Native either as their single race or as one of multiple races, regardless of their Hispanic origin.

- All non-Hispanic people not in Indian Country with American Indian or Alaska Native as their single race.

Domain 3 (Hispanic) includes:

- All Hispanic people in Indian Country, excluding those with American Indian or Alaska Native either as their single race or as one of multiple races.

- All Hispanic people not in Indian Country, excluding those who live in the state of Hawaii and have Native Hawaiian or Pacific Islander as a single race or as one of multiple races.

---

[1] Indian Country is land considered (either wholly or partially) on an American Indian reservation/trust land, Tribal Jurisdiction Statistical Area, Tribal Designated Statistical Area, or Alaska Native Village Statistical Area. For Census 2000, Tribal Jurisdiction Statistical Area has been formally renamed as Oklahoma Tribal Statistical Area.

Domain 4 (Non-Hispanic Black) includes:

- All non-Hispanic people with Black as their only race.

- All non-Hispanic people with the race combination of Black and American Indian or Alaska Native who do not live in Indian Country.

- All people with the race combination of Black and another single race group (Native Hawaiian or Pacific Islander, Asian, White, or "Some other race"), excluding those who live in the state of Hawaii and are Native Hawaiian or Pacific Islander in addition to Black.

Domain 5 (Native Hawaiian or Pacific Islander) includes:

- All non-Hispanic people with the single race Native Hawaiian or Pacific Islander.

- All non-Hispanic people with the race combination of Native Hawaiian or Pacific Islander and American Indian or Alaska Native who do not live in Indian Country.

- All non-Hispanic people with the race combination of Native Hawaiian or Pacific Islander and Asian.

- All people living in the state of Hawaii with Native Hawaiian or Pacific Islander race, regardless of their Hispanic origin and whether they identify with a single race or multiple races.

Domain 6 (Non-Hispanic Asian) includes:

- All non-Hispanic people with Asian as their single race.

- All people with the race combination of Asian and American Indian or Alaska Native who do not live in Indian Country.

Domain 7 (Non-Hispanic White or "Some other race") includes:

- All non-Hispanic people self-identifying as either White or "Some other race" as their single race, or self-identifying as both White and "Some other race."

- All non-Hispanic people with the race combination of American Indian or Alaska Native and White or "Some other race" who do not live in Indian Country.

- All non-Hispanic people with the race combinations of Asian and White or "Some other race."

- All non-Hispanic people with the race combination of Native Hawaiian or Pacific Islander and White or "Some other race," excluding those who live in the state of Hawaii.

- All non-Hispanic people with three or more races who live in Indian Country, excluding those with American Indian or Alaska Native as one of the races.

- All non-Hispanic people with three or more races and who do not live in Indian Country, excluding those who live in Hawaii and have Native Hawaiian or Pacific Islander as one of the races.

Table B1 shows the counts for the Race/Ethnicity domains. We used these counts as the denominators for the percent duplication estimates of the race/ethnicity domains. These counts are data-defined persons in housing units not including enumerations in Remote Alaska. Remote Alaska was out-of-scope for the A.C.E.

Table B1: Counts for Race/Ethnicity Domains

Race/Ethnicity Domain	Total
AI on AIR	513,147
AI off AIR	1,523,915
Hispanic	33,200,777
Non-Hispanic Black	32,330,425
Hawaiian and Pacific Islander	568,084
Non-Hispanic Asian	9,679,521
Non-Hispanic White or Some Other Race	190,004,235
Total	267,820,104

Table B2: Estimates of Nonmover Residents for Race/Ethnicity Domains

Race/Ethnicity Domain	Total[1]
AI on AIR	388,819
AI off AIR	1,172,984
Hispanic	30,561,120
Non-Hispanic Black	28,410,107
Hawaiian and Pacific Islander	528,746
Non-Hispanic Asian	8,556,605
Non-Hispanic White or Some Other Race	180,086,802
Total	249,705,183

[1] Total uses residence status determined for March, 2001 A.C.E. estimates

Appendix C: Age/Sex Categories

Table C1 shows the population counts for the Age/Sex categories. We used these counts as the denominators for the percent duplication estimates of the age/sex categories. These counts are data-defined persons in housing units not including enumerations in Remote Alaska. Remote Alaska was out-of-scope for the A.C.E.

Table C1: Counts for Age/Sex categories

Age/Sex Category	Total
0 - 9	38,289,359
10 - 17	31,419,609
18 - 29 Males	20,976,099
18 - 29 Females	21,024,109
30 - 49 Males	40,567,756
30 - 49 Females	42,105,085
50 + Males	33,375,084
50 + Females	40,063,003
Total	267,820,104

Table C2: Estimates of Nonmovers for Age/Sex categories

Age/Sex Category	Total[1]
0 - 17	65,808,126
18 - 29 Males	17,880,386
18 - 29 Females	18,369,923
30 - 49 Males	37,892,662
30 - 49 Females	40,725,014
50 + Males	31,117,004
50 + Females	37,912,069
Total	249,705,183

[1] Total used residence status determined for the March, 2001 estimates.

Appendix D: Assignment of the Unbiased Probability of Duplication (Multiplicity Weight)

E sample

For each duplication link between an E-sample record and a census enumeration identified by this analysis, we need to assign an unbiased probability of duplication or multiplicity weight. We can generate a design-based estimate of duplication based on this probability. Table D1 shows the combination of duplicates we estimated in this analysis.

<div align="center">

Table D1: Combinations of Duplicates

Combination
Duplication of E-sample Eligible to E-sample Eligible
Duplication of E-sample Eligible to GQ
Duplication of E-sample Eligible to Reinstate
Duplication of E-sample Eligible to Delete

</div>

Table D2 divides the records into 8 categories. The rest of this section describes how to assign probabilities based on the duplicate links between the categories.

<div align="center">

Table D2: Categories for Assigning Unbiased Probabilities

Category	Description
A	E-sample People in A.C.E. clusters
B	E-sample Eligible People not selected for the E sample (Both in the sample clusters and not in the sample clusters)
C	Group Quarters people in A.C.E. clusters
D	Group Quarters people not in A.C.E. clusters
E	Reinstated People in A.C.E. clusters
F	Reinstated People not in A.C.E. clusters
G	Deleted People in A.C.E. clusters
H	Deleted People not in A.C.E. clusters

</div>

Table D3 shows how to assign the unbiased probabilities. Each record represents a link between a E-sample person record and a census person record. The table shows how to the probability is assigned based on the type of link. When searching within the same universe (example: E-sample eligible to E-sample eligible), assigning a probability of 1 to each link would overestimate the amount of duplication. The table shows how to uses the number of links to

other records to assign an unbiased probability. This table lists only the combinations for the estimates in our analysis.

Table D3: Assignment of Unbiased Duplication Probabilities

E-sample Person to Census Link	Duplication Probability Value
A to A	$1\big/(U + 1)$
A to B	$\left(\dfrac{1}{U + V + 1}\right)\Big/(U + 1)$
A to C or D	$1\big/(U + 1)$
A to E or F	$1\big/(U + 1)$
A to G or H	$1\big/(U + 1)$

Where U is the number of links from this E-sample record to other category A records and V is the number of links from this E-sample record to category B records.

P Sample

The multiplicity weights for the P-sample links to the census enumerations are 1.

Since we are not matching within the same universe (E-sample eligible to E-sample Eligible) as we are on the E sample side, we don't need the factors that the E sample links to do.

Appendix E: Effect of Using Efficiency Measures Within the Cluster To Adjust Estimates of Duplication

Figure E1: One Person Household to One Person Household
Using Different Efficiency Adjustments

WC = Within Cluster SB = Surrounding Blocks SC = Same County SS = Different County, Same State DS = Different State

E-1

Figure E2: One Person Household to Two—Plus Person Household
Using Different Efficiency Adjustments
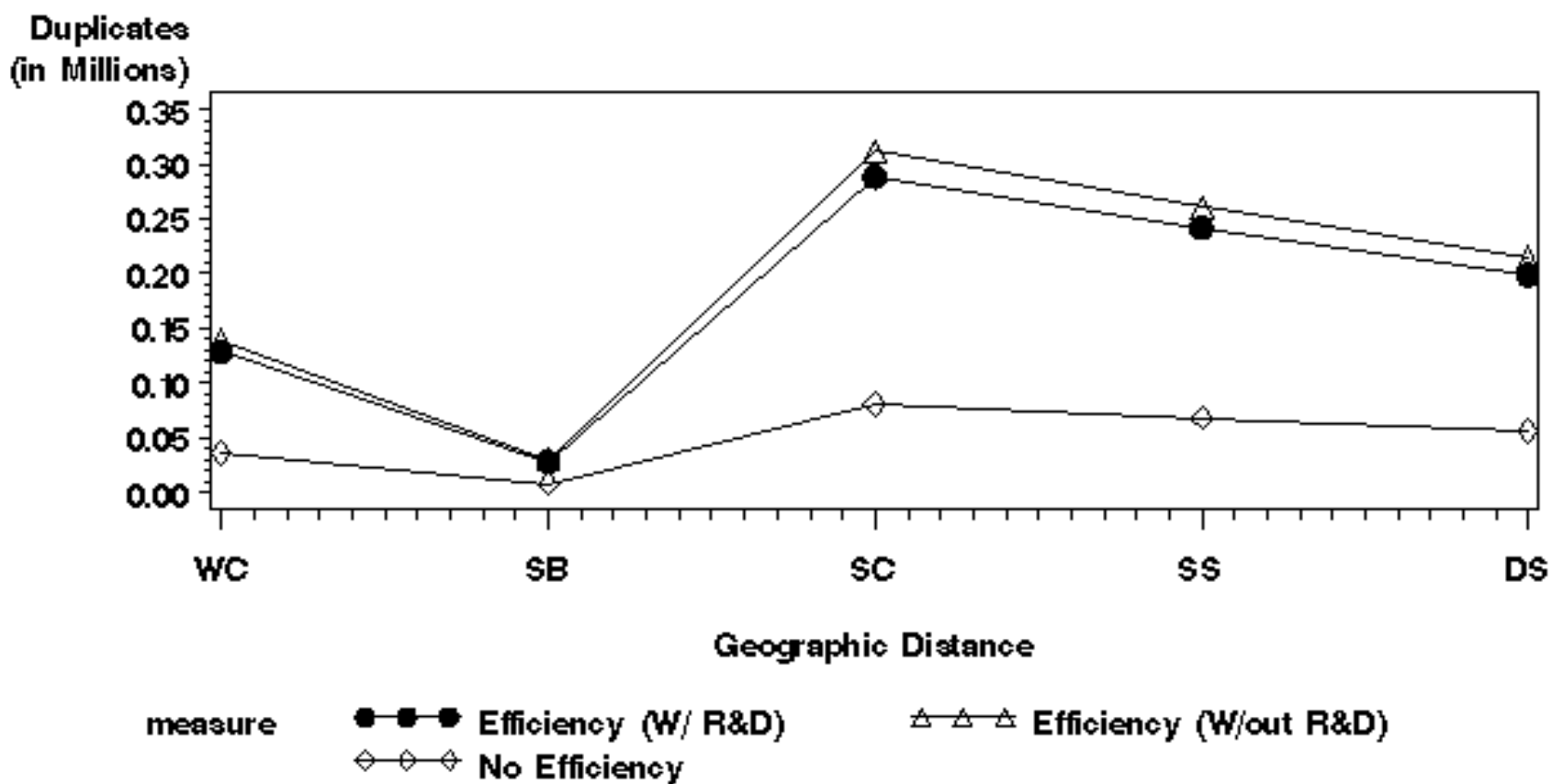
Figure E3: Whole HH Duplication Between Two—Plus Person Household
Using Different Efficiency Adjustments
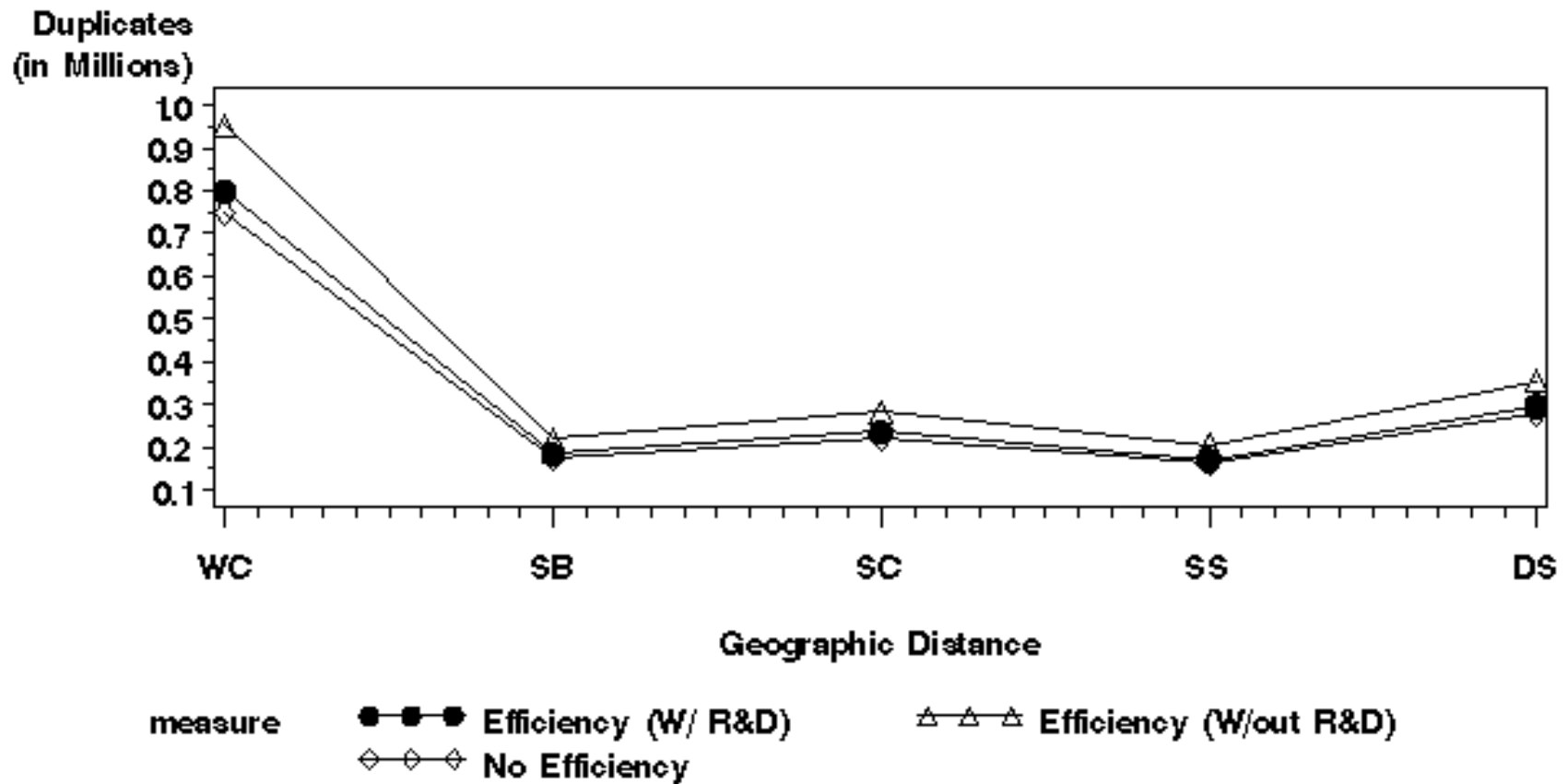
WC = Within Cluster   SB = Surrounding Blocks   SC = Same County   SS = Different County, Same State   DS = Different State

E-3

Figure E4: Partial Duplication Between Two—Plus Person Household
Using Different Efficiency Adjustments
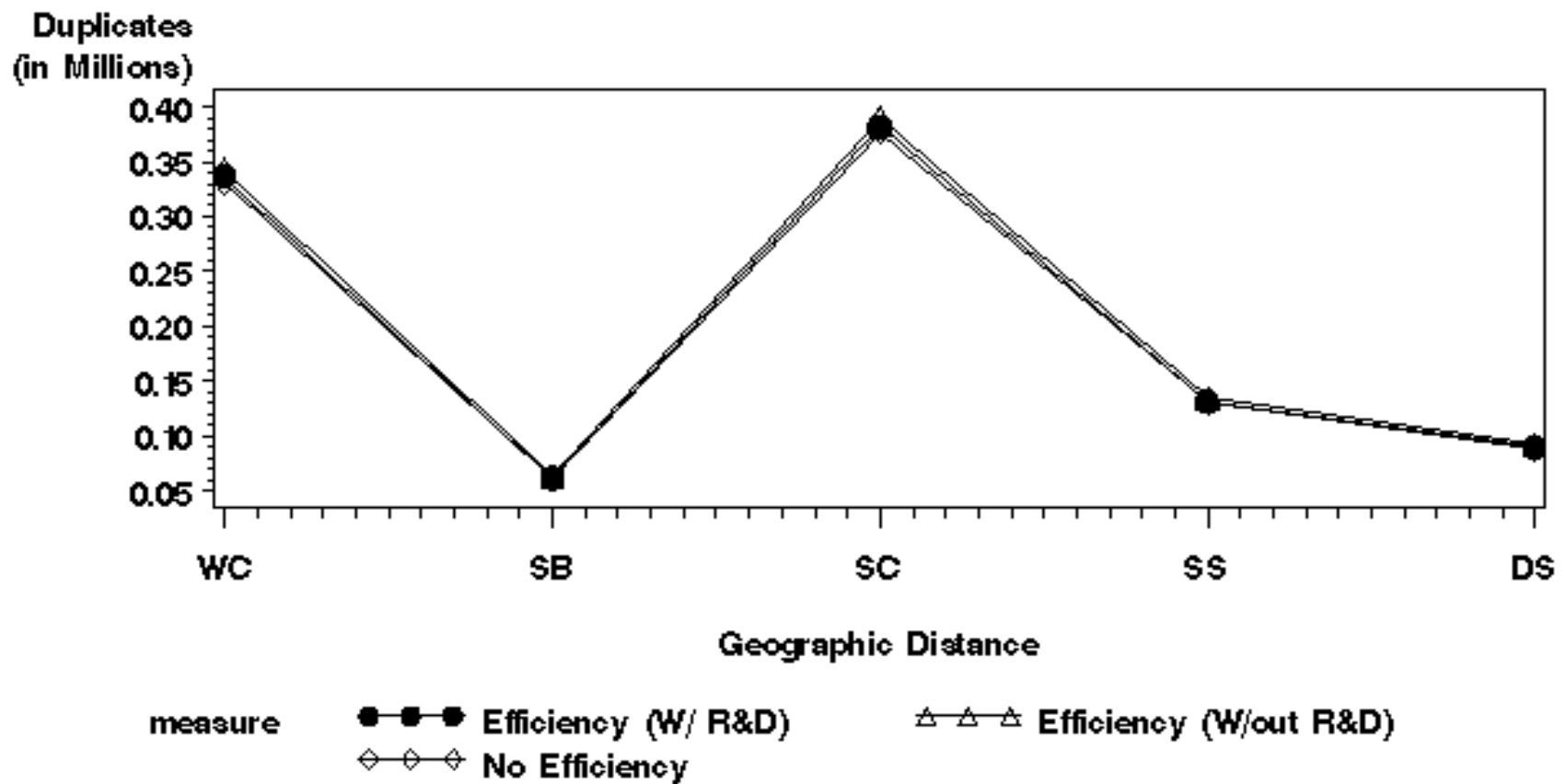
WC = Within Cluster SB = Surrounding Blocks SC = Same County SS = Different County, Same State DS = Different State

Figure E5: Only One Duplicate Between Two-Plus Person Household
Using Different Efficiency Adjustments
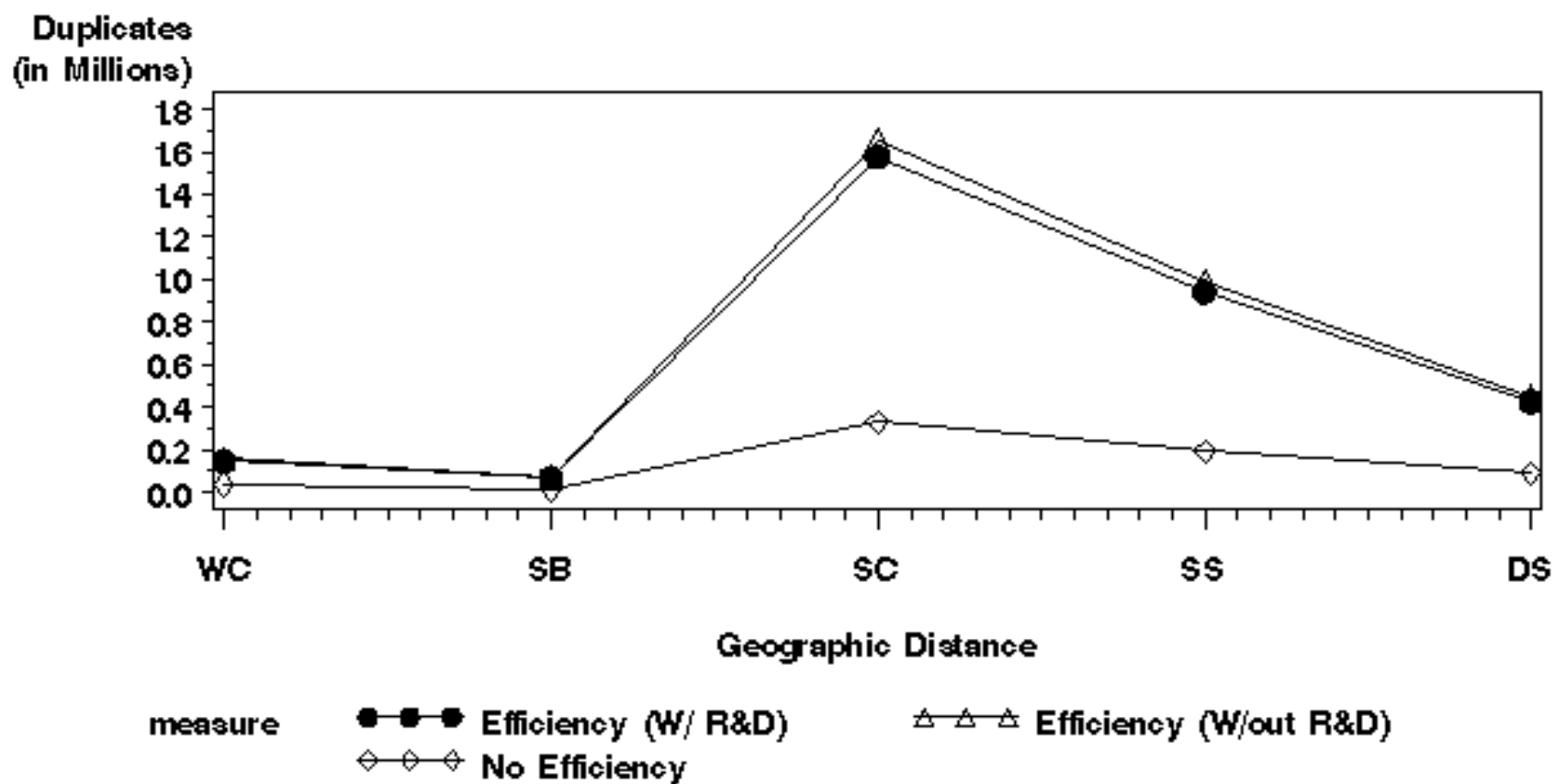
WC = Within Cluster   SB = Surrounding Blocks   SC = Same County   SS = Different County, Same State   DS = Different State

Table F1: Percent Duplication of Race/Ethnicity Domains by Geography
Census Housing Unit to Census Housing Unit (Total)[1]

Race/Ethnicity Domain	Total	Within Cluster[2]	Surrounding Blocks[2]	Outside Surrounding Blocks		
				Same County	Same State	Different State
Non-Hispanic White or Some Other Race	1.76% (0.04%)	0.69% (0.03%)	0.10% (0.01%)	0.49% (0.01%)	0.26% (0.01%)	0.22% (0.01%)
Non-Hispanic Black	2.46% (0.11%)	1.14% (0.08%)	0.13% (0.02%)	0.85% (0.05%)	0.18% (0.02%)	0.16% (0.02%)
Hispanic	2.43% (0.13%)	1.29% (0.09%)	0.12% (0.03%)	0.68% (0.04%)	0.19% (0.02%)	0.15% (0.06%)
Non-Hispanic Asian	2.08% (0.20%)	1.10% (0.17%)	0.12% (0.05%)	0.44% (0.06%)	0.24% (0.04%)	0.17% (0.03%)
Hawaiian and Pacific Islander	1.63% (0.53%)	0.57% (0.19%)	0.09% (0.06%)	0.88% (0.47%)	0.07% (0.03%)	0.01% (0.01%)
American Indian on AIR	2.74% (0.56%)	0.56% (0.13%)	0.41% (0.21%)	1.26% (0.36%)	0.26% (0.05%)	0.25% (0.05%)
American Indian off AIR	2.29% (0.41%)	0.74% (0.17%)	0.17% (0.12%)	0.71% (0.15%)	0.39% (0.14%)	0.28% (0.09%)

[1] Does not include duplicates where both were reinstated cases. These were included in the ESCAP II analysis. Not including these cases is negligible.

[2] This estimate is from our analysis and not A.C.E. Standard errors in parentheses.

Table F2: Percent Duplication of Race/Ethnicity Domains by Geography
Census Housing Unit to Census Housing Unit (Not Including Duplicates to Reinstated Units)

Race/ Ethnicity Domain	Total	Within Cluster[1]	Surrounding Blocks[1]	Outside Surrounding Blocks		
				Same County	Same State	Different State
Non-Hispanic White or Some Other Race	1.18% (0.02%)	0.33% (0.02%)	0.09% (0.01%)	0.32% (0.01%)	0.22% (0.01%)	0.22% (0.01%)
Non-Hispanic Black	1.76% (0.08%)	0.70% (0.06%)	0.11% (0.02%)	0.62% (0.03%)	0.16% (0.01%)	0.16% (0.02%)
Hispanic	1.65% (0.08%)	0.75% (0.06%)	0.12% (0.03%)	0.51% (0.03%)	0.18% (0.02%)	0.09% (0.01%)
Non-Hispanic Asian	1.38% (0.15%)	0.62% (0.12%)	0.12% (0.05%)	0.26% (0.04%)	0.20% (0.03%)	0.17% (0.03%)
Hawaiian and Pacific Islander	1.00% (0.24%)	0.53% (0.18%)	0.09% (0.06%)	0.30% (0.09%)	0.07% (0.03%)	0.01% (0.01%)
American Indian on AIR	2.47% (0.55%)	0.37% (0.10%)	0.40% (0.21%)	1.20% (0.36%)	0.25% (0.05%)	0.25% (0.05%)
American Indian off AIR	1.63% (0.29%)	0.40% (0.12%)	0.17% (0.12%)	0.48% (0.11%)	0.29% (0.09%)	0.28% (0.09%)

[1] This estimate is from our analysis and not A.C.E. Standard errors in parentheses.

Table F3: Percent Duplication of Race/Ethnicity Domains by Group Quarters Type
Census Housing Unit to Group Quarters

Race/ Ethnicity Domain	Total	Correctional Institution	Nursing Home	Juvenile Institution	College Dorm	Military	Other
Non-Hispanic White or Some Other Race	0.22% (0.01%)	0.02% (0.00%)	0.04% (0.00%)	0.00% (0.00%)	0.11% (0.01%)	0.01% (0.00%)	0.04% (0.00%)
Non-Hispanic Black	0.36% (0.02%)	0.11% (0.01%)	0.01% (0.00%)	0.02% (0.01%)	0.13% (0.01%)	0.01% (0.00%)	0.08% (0.01%)
Hispanic	0.16% (0.02%)	0.05% (0.01%)	0.01% (0.00%)	0.01% (0.00%)	0.04% (0.01%)	0.01% (0.00%)	0.04% (0.01%)
Non-Hispanic Asian	0.35% (0.16%)	0.00% (0.00%)	0.00% (0.00%)	0.00% (0.00%)	0.31% (0.16%)	0.00% (0.00%)	0.04% (0.01%)
Hawaiian and Pacific Islander	0.16% (0.07%)				0.06% (0.03%)		0.10% (0.06%)
American Indian on AIR	0.21% (0.04%)	0.05% (0.02%)	0.01% (0.01%)	0.04% (0.02%)	0.04% (0.02%)		0.06% (0.02%)
American Indian off AIR	0.20% (0.07%)	0.12% (0.05%)		0.03% (0.03%)	0.01% (0.01%)		0.04% (0.03%)

Standard errors in parentheses. Blanks indicate no duplicates were detected for that group.

Table F4: Percent Duplication of Race/Ethnicity Domains by Geography
Census Housing Unit to Housing Units Deleted During the HUDO in the Census

Race/ Ethnicity Domain	Total	Within Cluster	Surrounding Blocks	Outside Surrounding Blocks		
				Same County	Same State	Different State
Non-Hispanic White or Some Other Race	1.09% (0.04%)	0.72% (0.04%)	0.28% (0.03%)	0.07% (0.01%)	0.02% (0.00%)	0.01% (0.00%)
Non-Hispanic Black	1.02% (0.08%)	0.71% (0.07%)	0.21% (0.03%)	0.08% (0.02%)	0.01% (0.01%)	0.00% (0.00%)
Hispanic	1.14% (0.12%)	0.86% (0.11%)	0.17% (0.03%)	0.10% (0.02%)	0.01% (0.00%)	0.01% (0.01%)
Non-Hispanic Asian	1.12% (0.19%)	0.74% (0.13%)	0.25% (0.12%)	0.12% (0.05%)	0.00% (0.00%)	0.01% (0.01%)
Hawaiian and Pacific Islander	1.19% (0.38%)	0.53% (0.19%)	0.57% (0.31%)	0.09% (0.07%)		
American Indian on AIR	0.94% (0.30%)	0.56% (0.21%)	0.09% (0.05%)	0.29% (0.21%)		
American Indian off AIR	0.85% (0.20%)	0.52% (0.16%)	0.30% (0.12%)	0.01% (0.01%)		0.02% (0.02%)

Standard errors in parentheses. Blanks indicate no duplicates were detected for that group.

Table F5: Percent Duplication of Age/Sex Categories by Geography
Census Housing Unit to Census Housing Units (Total)[1]

Age/Sex Category	Total	Within Cluster[2]	Surrounding Blocks[2]	Outside Surrounding Blocks		
				Same County	Same State	Different State
0 - 9	2.33%	0.85%	0.11%	0.94%	0.26%	0.16%
	(0.06%)	(0.05%)	(0.01%)	(0.04%)	(0.02%)	(0.02%)
10 - 17	2.50%	0.91%	0.14%	1.03%	0.26%	0.16%
	(0.08%)	(0.05%)	(0.02%)	(0.04%)	(0.02%)	(0.04%)
18-29 Males	2.35%	0.93%	0.10%	0.75%	0.38%	0.19%
	(0.07%)	(0.05%)	(0.01%)	(0.03%)	(0.02%)	(0.01%)
18-29 Females	2.64%	0.99%	0.12%	0.88%	0.42%	0.23%
	(0.07%)	(0.06%)	(0.02%)	(0.04%)	(0.03%)	(0.02%)
30-49 Males	1.52%	0.80%	0.09%	0.36%	0.17%	0.10%
	(0.04%)	(0.03%)	(0.01%)	(0.02%)	(0.01%)	(0.01%)
30-49 Females	1.51%	0.80%	0.09%	0.36%	0.14%	0.11%
	(0.04%)	(0.03%)	(0.01%)	(0.02%)	(0.01%)	(0.01%)
50 + Males	1.81%	0.77%	0.11%	0.25%	0.28%	0.39%
	(0.05%)	(0.03%)	(0.01%)	(0.02%)	(0.02%)	(0.03%)
50 + Females	1.58%	0.74%	0.10%	0.21%	0.20%	0.33%
	(0.04%)	(0.03%)	(0.01%)	(0.01%)	(0.02%)	(0.02%)

[1] Does not include duplicates where both were reinstated cases. These were included in the ESCAP II analysis. Not including these cases is negligible.
[2] This estimate is from our analysis and not A.C.E. Standard errors in parentheses.

Table F6: Percent Duplication of Age/Sex Categories by Geography
Census Housing Unit to Census Housing Units (Not Including Duplicates to Reinstated Units)

Age/Sex Category	Total	Within Cluster[1]	Surrounding Blocks[1]	Outside Surrounding Blocks		
				Same County	Same State	Different State
0 - 9	1.53%	0.42%	0.10%	0.63%	0.22%	0.16%
	(0.04%)	(0.03%)	(0.01%)	(0.02%)	(0.01%)	(0.02%)
10 - 17	1.70%	0.45%	0.12%	0.77%	0.23%	0.13%
	(0.05%)	(0.03%)	(0.02%)	(0.03%)	(0.02%)	(0.01%)
18-29 Males	1.81%	0.60%	0.09%	0.58%	0.35%	0.19%
	(0.05%)	(0.04%)	(0.01%)	(0.03%)	(0.02%)	(0.01%)
18-29 Females	1.92%	0.56%	0.11%	0.64%	0.38%	0.23%
	(0.06%)	(0.04%)	(0.02%)	(0.03%)	(0.03%)	(0.02%)
30-49 Males	0.99%	0.44%	0.08%	0.23%	0.14%	0.09%
	(0.03%)	(0.02%)	(0.01%)	(0.01%)	(0.01%)	(0.01%)
30-49 Females	0.92%	0.41%	0.08%	0.21%	0.12%	0.10%
	(0.03%)	(0.02%)	(0.01%)	(0.01%)	(0.01%)	(0.01%)
50 + Males	1.25%	0.37%	0.11%	0.15%	0.23%	0.39%
	(0.04%)	(0.02%)	(0.01%)	(0.01%)	(0.01%)	(0.03%)
50 + Females	1.07%	0.38%	0.09%	0.11%	0.16%	0.32%
	(0.04%)	(0.02%)	(0.01%)	(0.01%)	(0.01%)	(0.02%)

[1] This estimate is from our analysis and not A.C.E. Standard errors in parentheses.

Table F7: Percent Duplication of Age/Sex Categories by Group Quarters Type
Census Housing Unit to Group Quarters

Age/Sex Category	Total	Correctional Institution	Nursing Home	Juvenile Institution	College Dorm	Military	Other
0 - 9	0.03% (0.01%)	0.00% (0.00%)		0.00% (0.00%)	0.00% (0.00%)		0.02% (0.01%)
10 - 17	0.11% (0.01%)	0.01% (0.00%)		0.06% (0.01%)	0.01% (0.00%)	0.00% (0.00%)	0.04% (0.01%)
18 - 29 Males	0.88% (0.08%)	0.19% (0.02%)	0.00% (0.00%)	0.00% (0.00%)	0.56% (0.07%)	0.06% (0.01%)	0.07% (0.01%)
18 - 29 Females	0.80% (0.07%)	0.02% (0.01%)			0.74% (0.07%)	0.01% (0.01%)	0.03% (0.01%)
30 - 49 Males	0.17% (0.02%)	0.09% (0.01%)	0.00% (0.00%)		0.02% (0.02%)	0.01% (0.00%)	0.05% (0.01%)
30 - 49 Females	0.06% (0.01%)	0.02% (0.00%)	0.00% (0.00%)		0.01% (0.01%)		0.03% (0.01%)
50+ Males	0.15% (0.01%)	0.01% (0.00%)	0.08% (0.01%)		0.00% (0.00%)		0.06% (0.01%)
50+ Females	0.18% (0.02%)	0.00% (0.00%)	0.12% (0.02%)		0.00% (0.00%)		0.06% (0.01%)

Standard errors in parentheses. Blanks indicate no duplicates were detected for that group.

Table F8: Percent of Person Duplication of Age/Sex Categories
Census Housing Unit to Housing Units Deleted During the HUDO in the Census

Age/Sex Category	Total	Within Cluster	Surrounding Blocks	Outside Surrounding Blocks		
				Same County	Same State	Different State
0 - 9	1.19%	0.80%	0.27%	0.11%	0.02%	0.00%
	(0.06%)	(0.05%)	(0.04%)	(0.01%)	(0.01%)	(0.00%)
10 - 17	1.19%	0.81%	0.24%	0.11%	0.02%	0.01%
	(0.06%)	(0.05%)	(0.03%)	(0.02%)	(0.01%)	(0.01%)
18-29 Males	1.01%	0.73%	0.18%	0.07%	0.01%	0.01%
	(0.07%)	(0.06%)	(0.03%)	(0.02%)	(0.01%)	(0.00%)
18-29 Females	1.22%	0.87%	0.24%	0.09%	0.01%	0.01%
	(0.07%)	(0.06%)	(0.03%)	(0.01%)	(0.01%)	(0.00%)
30-49 Males	1.04%	0.71%	0.25%	0.07%	0.01%	0.00%
	(0.05%)	(0.04%)	(0.03%)	(0.01%)	(0.00%)	(0.00%)
30-49 Females	1.06%	0.73%	0.25%	0.07%	0.01%	0.00%
	(0.04%)	(0.04%)	(0.02%)	(0.01%)	(0.00%)	(0.00%)
50 + Males	1.06%	0.69%	0.29%	0.05%	0.01%	0.01%
	(0.05%)	(0.04%)	(0.04%)	(0.01%)	(0.00%)	(0.00%)
50 + Females	0.97%	0.62%	0.26%	0.06%	0.01%	0.01%
	(0.04%)	(0.03%)	(0.03%)	(0.01%)	(0.00%)	(0.00%)

Table F9: Person Duplication Results by Tenure
Census Housing Unit to Census Housing Unit (Total)[1]

	Total	Within Cluster[2]	Surrounding Block[2]	Outside Surrounding Blocks		
				Same County	Different County, Same State	Different State
Both Owners	2,471,016 (60,251)	1,161,278 (49,010)	169,706 (16,101)	529,414 (21,360)	316,562 (15,872)	294,056 (18,718)
One Owner/ One Renter	1,253,171 (34,303)	220,391 (19,404)	41,924 (5,932)	567,758 (20,423)	248,705 (12,919)	174,393 (13,085)
Both Renters	1,486,552 (49,287)	850,224 (40,199)	72,927 (9,723)	396,763 (19,935)	86,333 (7,865)	80,305 (12,741)

[1] Does not include duplicates where both were reinstated cases. These were included in the ESCAP II analysis. Not including these cases is negligible.
[2] This estimate is from our analysis and not A.C.E. Standard errors in parentheses.

Table F10: Person Duplication Results by Tenure
Census Housing Unit to Census Housing Units (Not Including Duplicates to Reinstated Units)

| | Total | Geography | | | | |
| | | Within Cluster[1] | Surrounding Block[1] | Outside Surrounding Blocks | | |
				Same County	Different County, Same State	Different State
Both Owners	1,584,111 (40,972)	531,492 (28,855)	154,092 (15,029)	345,410 (12,803)	261,173 (11,943)	291,944 (18,661)
One Owner/ One Renter	1,017,779 (24,925)	153,876 (14,048)	38,210 (5,262)	432,417 (14,755)	227,778 (11,291)	165,498 (8,765)
Both Renters	934,245 (34,665)	487,977 (28,556)	67,502 (9,315)	234,094 (11,264)	74,319 (6,855)	70,354 (8,514)

[1] This estimate is from our analysis and not A.C.E. Standard errors in parentheses.

Table F11: Person Duplication Results by Tenure
Census Housing Unit to Group Quarters

| | Total | Type of Group Quarters | | | | | |
		Correctional Institution	Nursing Home	Juvenile Institution	College Dorms	Military	Other
Owner	338,172 (10,544)	39,011 (5,079)	39,731 (4,816)	8,835 (2,074)	192,659 (8,184)	12,591 (2,538)	45,346 (6,007)
Renter	277,566 (42,735)	53,905 (5,638)	34,871 (6,954)	10,042 (2,025)	97,534 (41,516)	7,932 (1,873)	73,281 (8,765)

Standard errors in parentheses.

Table F12: Person Duplication Results by Tenure
Census Housing Unit to Persons Deleted in HUDO in the Census

	Total	Geography				
		Within Cluster	Surrounding Block	Outside Surrounding Blocks		
				Same County	Different County, Same State	Different State
Both Owners	1,847,662 (84,302)	1,225,260 (67,824)	506,207 (50,431)	94,866 (13,927)	15,624 (5,272)	5,706 (2,315)
One Owner/ One Renter	227,656 (19,490)	130,702 (14,466)	39,570 (8,937)	36,379 (8,017)	11,807 (3,849)	9,199 (4,239)
Both Renters	829,777 (48,012)	611,236 (42,893)	132,579 (17,439)	77,002 (12,310)	7,680 (3,143)	1280 (753)

Standard errors in parentheses.

Table G1: Percent of FSPD Matches of P-sample Nonmovers by Race/Ethnicity Domains and Geography
P-sample Nonmover Residents to Census Housing Unit (Total)

Race/Ethnicity Domain	Total	Within Cluster[1]	Surrounding Blocks[1]	Outside Surrounding Blocks		
				Same County	Same State	Different State
Non-Hispanic White or Some Other Race	88.00% (0.16%)	82.27% (0.24%)	3.64% (0.19%)	1.21% (0.05%)	0.44% (0.02%)	0.43% (0.01%)
Non-Hispanic Black	80.66% (0.37%)	73.64% (0.55%)	4.23% (0.48%)	2.03% (0.11%)	0.36% (0.03%)	0.40% (0.03%)
Hispanic	84.57% (0.42%)	77.27% (0.60%)	4.36% (0.59%)	1.79% (0.13%)	0.50% (0.03%)	0.64% (0.04%)
Non-Hispanic Asian	82.20% (0.63%)	76.49% (0.68%)	3.82% (0.69%)	1.23% (0.19%)	0.38% (0.05%)	0.28% (0.04%)
Hawaiian and Pacific Islander	79.81% (1.40%)	71.53% (1.56%)	5.24% (1.90%)	2.82% (1.55%)	0.19% (0.13%)	0.02% (0.02%)
American Indian on AIR	87.24% (0.94%)	79.17% (1.51%)	3.85% (0.91%)	3.18% (0.89%)	0.66% (0.11%)	0.37% (0.06%)
American Indian off AIR	84.88% (1.17%)	77.16% (1.46%)	4.31% (1.46%)	1.63% (0.38%)	1.35% (0.42%)	0.42% (0.11%)

[1] This estimate is from our analysis and not A.C.E.
Note: These estimates include the residence probability. We used the residence probability from the March, 2001 estimates. Standard errors in parentheses.

Table G2: Percent of FSPD Matches of P-sample Nonmovers by Race/Ethnicity Domains and Geography
P-sample Nonmover Residents to Census Housing Units (Not Including Duplicates to Reinstated Units)

Race/Ethnicity Domain	Total	Within Cluster[1]	Surrounding Blocks[1]	Outside		
				Same County	Same State	Different State
Non-Hispanic White or Some Other Race	87.34% (0.16%)	81.76% (0.24%)	3.60% (0.19%)	1.13% (0.05%)	0.42% (0.02%)	0.43% (0.01%)
Non-Hispanic Black	79.79% (0.37%)	72.97% (0.56%)	4.16% (0.48%)	1.91% (0.11%)	0.36% (0.03%)	0.40% (0.03%)
Hispanic	83.77% (0.41%)	76.59% (0.60%)	4.32% (0.59%)	1.73% (0.13%)	0.50% (0.03%)	0.63% (0.04%)
Non-Hispanic Asian	81.32% (0.65%)	75.72% (0.73%)	3.82% (0.69%)	1.18% (0.19%)	0.33% (0.05%)	0.28% (0.04%)
Hawaiian and Pacific Islander	78.65% (1.29%)	71.07% (1.57%)	5.15% (1.90%)	2.23% (1.48%)	0.19% (0.13%)	0.02% (0.02%)
American Indian on AIR	86.72% (0.95%)	78.89% (1.51%)	3.72% (0.89%)	3.09% (0.89%)	0.65% (0.12%)	0.37% (0.06%)
American Indian off AIR	84.21% (1.10%)	76.64% (1.43%)	4.31% (1.46%)	1.58% (0.38%)	1.25% (0.38%)	0.42% (0.11%)

[1] This estimate is from our analysis and not A.C.E.
Note: These estimates include the residence probability. We used the residence probability from the March, 2001 estimates. Standard errors in parentheses.

Table G3: Percent of FSPD Matches of P-sample Nonmovers by Race/Ethnicity Domains and Type of Group Quarters
P-sample Nonmover Residents to Group Quarters

Race/ Ethnicity Domain	Total	Correctional Institution	Nursing Home	Juvenile Institution	College Dorm	Military	Other
Non-Hispanic White or Some Other Race	0.16% (0.01%)	0.03% (0.00%)	0.04% (0.01%)	0.00% (0.00%)	0.04% (0.01%)	0.01% (0.00%)	0.05% (0.01%)
Non-Hispanic Black	0.40% (0.03%)	0.11% (0.01%)	0.02% (0.01%)	0.04% (0.01%)	0.08% (0.01%)	0.02% (0.01%)	0.12% (0.02%)
Hispanic	0.24% (0.02%)	0.08% (0.01%)	0.01% (0.00%)	0.02% (0.01%)	0.04% (0.01%)	0.01% (0.00%)	0.07% (0.01%)
Non-Hispanic Asian	0.27% (0.12%)	0.01% (0.01%)	0.03% (0.02%)		0.21% (0.12%)	0.01% (0.00%)	0.02% (0.01%)
Hawaiian and Pacific Islander	0.37% (0.16%)	0.02% (0.02%)			0.04% (0.03%)		0.31% (0.16%)
American Indian on AIR	0.28% (0.05%)	0.09% (0.03%)	0.00% (0.00%)	0.03% (0.02%)	0.06% (0.03%)		0.09% (0.03%)
American Indian off AIR	0.14% (0.06%)	0.09% (0.05%)		0.00% (0.00%)	0.04% (0.03%)		0.01% (0.01%)

Note: These estimates include the residence probability of the nonmover case. We used the residence probability from the March, 2001 estimates. Standard errors in parentheses. Blanks indicates no matches were detected for that group.

Table G4: Percent of FSPD Matches of P-sample Nonmovers by Race/Ethnicity Domains and Geography
P-sample Nonmover Residents to Housing Units Deleted During HUDO in the Census

Race/Ethnicity Domain	Total	Within Cluster[1]	Surrounding Blocks[1]	Outside Surrounding Blocks		
				Same County	Same State	Different State
Non-Hispanic White or Some Other Race	1.11% (0.06%)	0.91% (0.06%)	0.14% (0.01%)	0.05% (0.01%)	0.01% (0.00%)	0.01% (0.00%)
Non-Hispanic Black	1.16% (0.12%)	0.90% (0.11%)	0.16% (0.03%)	0.07% (0.02%)	0.01% (0.01%)	0.00% (0.00%)
Hispanic	1.25% (0.15%)	1.02% (0.13%)	0.10% (0.02%)	0.11% (0.07%)	0.01% (0.01%)	0.01% (0.00%)
Non-Hispanic Asian	1.05% (0.19%)	0.78% (0.13%)	0.18% (0.11%)	0.10% (0.05%)		0.00% (0.00%)
Hawaiian and Pacific Islander	0.75% (0.25%)	0.51% (0.21%)	0.22% (0.16%)	0.02% (0.02%)		
American Indian on AIR	1.63% (0.79%)	1.09% (0.42%)	0.09% (0.06%)	0.46% (0.43%)		
American Indian off AIR	1.79% (0.53%)	1.50% (0.53%)	0.29% (0.14%)			0.01% (0.01%)

Note: These estimates include the residence probability of the nonmover case. We used the residence probability from the March, 2001 estimates. Standard errors in parentheses. Blanks indicate no matches were detected for this group.

Table G5: Percent of FSPD Matches of P-sample Nonmovers by Age/Sex Categories and Geography
P-sample Nonmover Residents to Census Housing Unit (Total)

Age/Sex Category	Total	Within Cluster[1]	Surrounding Blocks[1]	Outside Surrounding Blocks		
				Same County	Same State	Different State
0 - 17	91.80% (0.18%)	84.55% (0.34%)	4.18% (0.31%)	2.21% (0.08%)	0.49% (0.02%)	0.37% (0.02%)
18-29 Males	81.62% (0.32%)	74.79% (0.38%)	3.67% (0.28%)	1.88% (0.09%)	0.79% (0.05%)	0.48% (0.03%)
18-29 Females	85.79% (0.27%)	78.42% (0.37%)	4.07% (0.30%)	2.10% (0.09%)	0.70% (0.05%)	0.50% (0.03%)
30-49 Males	84.09% (0.22%)	78.74% (0.30%)	3.56% (0.21%)	1.07% (0.06%)	0.35% (0.02%)	0.37% (0.02%)
30-49 Females	87.44% (0.19%)	81.99% (0.28%)	3.86% (0.23%)	0.97% (0.05%)	0.30% (0.02%)	0.31% (0.02%)
50 + Males	85.25% (0.23%)	79.81% (0.28%)	3.61% (0.21%)	0.76% (0.05%)	0.40% (0.03%)	0.67% (0.03%)
50 + Females	82.43% (0.24%)	77.40% (0.30%)	3.46% (0.22%)	0.66% (0.05%)	0.32% (0.02%)	0.58% (0.03%)

[1] This estimate is from our analysis and not A.C.E.
Note: These estimates include the residence probability of the nonmover case. We used the residence probability from the March, 2001 estimates. Standard errors in parentheses.

Table G6: Percent of FSPD Matches of P-sample Nonmovers by Age/Sex Categories and Geography
P-sample Nonmover Residents to Census Housing Units (Not Including Duplicates to Reinstated Units)

Age/Sex Category	Total	Within Cluster[1]	Surrounding Blocks[1]	Outside Surrounding Blocks		
				Same County	Same State	Different State
0 - 17	90.96%	83.93%	4.11%	2.07%	0.48%	0.37%
	(0.17%)	(0.34%)	(0.31%)	(0.08%)	(0.02%)	(0.02%)
18-29 Males	80.97%	74.26%	3.65%	1.81%	0.77%	0.48%
	(0.32%)	(0.38%)	(0.28%)	(0.09%)	(0.05%)	(0.03%)
18-29 Females	84.98%	77.80%	4.02%	1.98%	0.68%	0.50%
	(0.26%)	(0.37%)	(0.30%)	(0.09%)	(0.05%)	(0.03%)
30-49 Males	83.49%	78.25%	3.52%	1.01%	0.34%	0.37%
	(0.22%)	(0.30%)	(0.21%)	(0.06%)	(0.02%)	(0.02%)
30-49 Females	86.77%	81.46%	3.81%	0.91%	0.29%	0.31%
	(0.19%)	(0.28%)	(0.23%)	(0.05%)	(0.02%)	(0.02%)
50 + Males	84.62%	79.29%	3.57%	0.71%	0.38%	0.67%
	(0.23%)	(0.28%)	(0.21%)	(0.05%)	(0.02%)	(0.03%)
50 + Females	81.80%	76.87%	3.43%	0.61%	0.30%	0.58%
	(0.24%)	(0.30%)	(0.21%)	(0.05%)	(0.02%)	(0.03%)

[1] This estimate is from our analysis and not A.C.E.

Note: These estimates include the residence probability of the nonmover case. We used the residence probability from the March, 2001 estimates. Standard errors in parentheses.

Table G7: Percent of FSPD Matches of P-sample Nonmovers by Age/Sex Categories and Type of Group Quarters
P-sample Nonmover Residents to Group Quarters

Age/Sex Category	Total	Correctional Institution	Nursing Home	Juvenile Institution	College Dorm	Military	Other
0 - 17	0.07%	0.00%		0.03%	0.00%		0.03%
	(0.01%)	(0.00%)		(0.00%)	(0.00%)		(0.01%)
18-29 Males	0.67%	0.22%	0.00%	0.01%	0.28%	0.08%	0.07%
	(0.06%)	(0.02%)	(0.00%)	(0.01%)	(0.05%)	(0.01%)	(0.01%)
18-29 Females	0.45%	0.02%	0.00%		0.32%	0.01%	0.08%
	(0.05%)	(0.01%)	(0.00%)		(0.05%)	(0.01%)	(0.02%)
30-49 Males	0.24%	0.12%	0.00%		0.01%	0.02%	0.08%
	(0.02%)	(0.01%)	(0.00%)		(0.01%)	(0.01%)	(0.01%)
30-49 Females	0.07%	0.01%	0.00%	0.00%	0.01%	0.00%	0.04%
	(0.01%)	(0.00%)	(0.00%)	(0.00%)	(0.01%)	(0.00%)	(0.01%)
50 + Males	0.19%	0.03%	0.08%		0.00%		0.08%
	(0.03%)	(0.01%)	(0.02%)		(0.00%)		(0.02%)
50 + Females	0.19%	0.00%	0.12%		0.00%		0.07%
	(0.04%)	(0.00%)	(0.04%)		(0.00%)		(0.01%)

Note: These estimates include the residence probability of the nonmover case. We used the residence probability from the March, 2001 estimates. Standard errors in parentheses. Blanks indicate no matches were detected for this group.

Table G8: Percent of FSPD Matches of P-sample Nonmovers by Age/Sex Categories and Geography
P-sample Nonmover Residents to Housing Units Deleted During HUDO in the Census

Age/Sex Category	Total	Within Cluster	Surrounding Blocks	Outside Surrounding Blocks		
				Same County	Same State	Different State
0 - 17	1.21%	0.97%	0.13%	0.09%	0.01%	0.01%
	(0.07%)	(0.06%)	(0.02%)	(0.02%)	(0.00%)	(0.00%)
18-29 Males	1.02%	0.85%	0.10%	0.05%	0.01%	0.01%
	(0.09%)	(0.09%)	(0.02%)	(0.01%)	(0.00%)	(0.00%)
18-29 Females	1.21%	0.98%	0.12%	0.08%	0.02%	0.01%
	(0.08%)	(0.08%)	(0.02%)	(0.02%)	(0.01%)	(0.00%)
30-49 Males	1.02%	0.82%	0.13%	0.07%	0.00%	0.00%
	(0.06%)	(0.05%)	(0.02%)	(0.02%)	(0.00%)	(0.00%)
30-49 Females	1.11%	0.91%	0.13%	0.06%	0.01%	0.00%
	(0.07%)	(0.06%)	(0.02%)	(0.01%)	(0.00%)	(0.00%)
50 + Males	1.19%	0.96%	0.19%	0.03%	0.01%	0.01%
	(0.08%)	(0.08%)	(0.02%)	(0.01%)	(0.00%)	(0.00%)
50 + Females	1.10%	0.90%	0.16%	0.03%	0.00%	0.01%
	(0.08%)	(0.08%)	(0.02%)	(0.01%)	(0.00%)	(0.00%)

Note: These estimates include the residence probability of the nonmover case. We used the residence probability from the March, 2001 estimates. Standard errors in parentheses.