

# THE 1990 HOUSING UNIT COVERAGE STUDY

Danny R. Childers\*  
Bureau of the Census, Washington, DC 20233

**Key words:** housing coverage, census geography, undercount, coverage evaluation

## 1.0 Background

The Housing Unit Coverage Study (HUCS) was designed to produce an estimate of the net under-coverage of housing units in the form of a dual system estimate. The HUCS sample consisted of half of the sample from the Post Enumeration Survey (PES). The HUCS sample consisted of two parts, the P-sample and the E-sample. The P-sample was an independent listing of housing units in the sample blocks. The E-sample was the housing units enumerated in the same sample of blocks. The P-sample and the E-sample were overlapping samples. The P-sample estimated the gross percentage of housing units not matched to the census and the E-sample estimated the gross percentage of housing units erroneously enumerated in the census.

The addresses for the housing units in the HUCS sample were computer matched. The match results were reviewed clerically. Housing units that did not match, unresolved cases, and possible duplicates were sent to the field for a follow-up interview. The results of the follow-up interview were recorded and the final match results tabulated.

This design for HUCS required that the P-sample and the E-sample of housing units balance. One of the improvements of the 1990 design over the 1980 design was that balancing was assured, because the two samples overlapped and the search area was fixed. In 1980 the P-sample was from the Current Population Survey (CPS) and the E-sample was selected independently from the enumerated census housing units. The search area was fixed for the E-sample, but not for the P-sample, which reduced the reliability of the estimates. The search area was the land area where searching for a P-sample housing unit match in the census was allowed. When the E-sample housing unit enumerated within the sample block was actually located outside the search area, the housing unit was classified as a geographical coding error or geocode error.

Balancing in the two samples was required because of geocoding error in the census. A P-sample nonmatch means the housing unit was not located within the search area. It may have been counted in the census, but not in the search area. The housing unit may have been incorrectly geocoded somewhere outside the search area, making matching operationally impossible. Even though the P-sample nonmatches are not necessarily missed housing units, studying trends for various variables is instructive. We need to learn more about the kinds of nonmatches and where they occur in the 1990 Decennial Census in order to improve census methodology for the 2000 Decennial Census. A discussion of the P-sample nonmatches is presented in Section 2.0.

It is equally important to study the erroneous enumerations in the 1990 Decennial Census to prepare for the next census. A discussion of the E-sample erroneous enumerations is in Section 3.0. This section includes an investigation of the erroneous enumerations by several variables, such as type of enumeration area, occupancy status, and number of units in structure. This section also describes the four different classifications of erroneous enumeration, which are within block duplicate, nonexistent, not a housing unit, and geocoding error.

The Census Bureau is considering using the existing list of addresses as the base of the 2000 address list. In 1990 the address lists were created from purchasing mailing lists and

compiling address lists. Before the Bureau uses the 1990 list of addresses as a base for the 2000 census, the existing address list or address control file (ACF) should be evaluated. We can investigate the completeness of the existing addresses of the 1990 housing units by investigating the correct enumerations from HUCS. The quality of the address and the source of the information used to make a match are described in Section 4.0.

The quality of the geographic assignment of the housing units in the census address files is evaluated by estimating the percentage of housing units incorrectly geocoded to the census block. If the 1990 ACF is to be used for the 2000 census, the quality of census geocoding should be investigated. A discussion of the quality of census geocoding is in Section 5.0.

The method of estimating the 1990 coverage of housing units is different from 1980. The percentage not matched in both censuses and a discussion of both evaluations are presented in Section 6.0.

## 2.0 The P-Sample

The P-sample housing units were coded as matched, not matched, or unresolved after the follow-up results were reviewed. The estimates of the percentage in each category weighted to national totals are in Table 2.1. This study estimated the gross percentage of the housing units not matched in the census to be 3.4 percent. A housing unit was coded matched when the housing unit was classified as enumerated in the search area and not matched when the housing unit could not be located within the search area.

A housing unit was coded unresolved when the field interview for HUCS was inconclusive. The probability that the housing unit was matched to the census will be imputed for the unresolved cases in order to calculate dual system estimates. A logistic regression model will be fit to the P-sample data for which enumeration status was observed. This model will be used to predict the probability of correctly enumerated versus that of not matched to the Census for unresolved P-sample housing units.

**Table 2.1: P-sample Results Before Imputation**

Matched	96.0
Not Matched	3.4
Unresolved	0.6
Total	100.0

In order to analyze the 1990 Census of Population and Housing before we prepare for the next census, the universe not matched was investigated looking at a number of variables. The first was type of enumeration area. The method of census taking was different for the four types of enumeration areas: Tape Address Register (TAR), prelist, update/leave (U/L), and list/enumerate (L/E).

The TAR areas were inner city, urban, and some suburban areas. The census was conducted in TAR areas by purchasing mailing lists and after checking and updating the lists, the questionnaires in TAR areas were mailed. The TAR method of enumeration covered 56 percent of the population and 56 percent of the housing units in the United States.

The prelist areas were suburban and more rural areas. Address lists were developed by a field listing procedure in prelist areas followed by postal and field checks and as in TAR areas, questionnaires were mailed. The prelist method of enumeration covered 30 percent of the population and 29 percent of the housing units.

The update/leave areas were mainly Appalachia and the rural south. The postal service did not assist in developing the address list nor in delivery in update/leave areas. Addresses were field listed and updated during questionnaire delivery by Census staff. The update/leave method of enumeration covered 10 percent of the population and 10 percent of the housing units.

The list/enumerate areas represented the most rural sections of the country. Census staff prepared the list of addresses in list/enumerate areas as they enumerated the households. The list/enumerate method of enumeration covered 4.5 percent of the US population and 5.6 percent of the housing units.

The search areas for the different methods of census enumeration were not the same. The search area was the land area where searching for a P-sample housing unit match in the census was allowed. This search area was the sample block and the surrounding blocks. In TAR areas, the surrounding blocks consisted of one ring of blocks surrounding the sample block at any one point. For prelist and update/leave areas, the search area was the sample block and two rings of blocks surrounding the sample block. For list/enumerate areas, the search area was the entire address register area (ARA). The search area was designed to be smaller in the urban areas where we purchased mailing lists, because geocoding error was expected to be less.

The estimated percentage not matched is in Table 2.2 for the four types of enumeration areas. Estimated standard errors for the estimates of percentage not matched are in parentheses. Since the percentage not matched was not the same for the four types of enumeration areas, the percent not matched can not be compared for all types of enumeration areas. The percent not matched for prelist and update/leave can be compared, since the search areas were the same. The percentage not matched for prelist was 4.9 percent and for update/leave was 5.0 percent. These estimates of percent not matched were not significantly different.

**Table 2.2: Percentage Not Matched by Type of Enumeration Area**

TAR	2.1	(0.2)
Prelist	4.9	(0.5)
Update/Leave	5.0	(0.4)
List/Enumerate	6.4	(0.9)
Total	3.4	(0.2)

A comparison of the percentage not matched in occupied and vacant housing units is in Table 2.3. Vacant housing units were estimated to have a larger percentage not matched than occupied housing units. One reason for this may have been the problem with units that were vacant-boarded up and unfit for habitation. A housing unit that was boarded up should have been classified as fit for habitation, if the housing definition was met. The percentage not matched for vacant housing units was significantly greater than for occupied housing units.

Some people felt the under-enumerations in occupied housing units were more important, because the missed vacant housing units did not add to the undercount of persons. On the other hand, if the Bureau uses the existing file of census addresses as the base for the 2000 Decennial Census, the missed vacant housing units could be occupied in the year 2000. It should be noted that the occupancy status represented the status of the housing unit at the time of the PES interview, which was in July or August of 1990.

**Table 2.3: Percentage Not Matched by Occupancy Status of PES Interview**

Occupied	2.4	(0.2)
Vacant	11.4	(1.0)
Total	3.4	(0.2)

The estimated percentage not matched is in Table 2.4 for the four census regions. The estimated percentage not matched for the South was significantly different from the Midwest.

**Table 2.4: Percentage Not Matched by Census Region**

Northeast	4.1	(0.5)
South	3.9	(0.2)
Midwest	2.6	(0.4)
West	2.8	(0.5)

The estimated percentage not matched is presented in Table 2.4 for the five categories of number of units in structure. The category of "Other" included the trailers and any other type of housing unit that was not a structure, such as tents, van, and boats. The estimated percentage not matched for housing units in 2 to 9 unit structures was 5.6 percent. The high percentage not matched in this group may be housing unit conversions not included in the census file of addresses. The estimated percentage not matched in single units was 3.0 percent.

The estimated percentage not matched in housing units classified as single was significantly different from units in 2 to 9 unit structures. The estimated percentage not matched in housing units classified as "other" was significantly different from the single housing units.

The estimated percentage not matched for all multi-unit structures was 3.5 percent with a standard error of 0.4 percent. The estimated percentage not matched for all multi-unit structures was not significantly different from single unit structures, but was significantly different from the structures classified as "other".

**Table 2.5: Percentage Not Matched by Number of Units in Structure**

Single	3.0	(0.2)
2 to 9	5.6	(0.5)
10 to 49	1.6	(0.4)
50 and over	1.3	(0.8)
Other	7.1	(1.1)

The estimated percentage not matched is in Table 2.6 for the five race/origin categories for occupied housing units. The race/origin classification was obtained from the first person in the household for the occupied P-sample housing units. The estimated percentage not matched for housing units classified as "Black" was not significantly different from the housing units classified as "Hispanic". The estimated percentage not matched for housing units classified as "Black" was not significantly different from housing units classified as "White and Other".

The estimate of the percentage not matched for housing units classified as American Indian was 5.2 percent. Some of these housing units were on reservations. These housing units may have been not matched to the census because of the omission in the census of non-traditional types of living situations on the reservations.

**Table 2.6: Percentage Not Matched by Race/Origin of Person 1**

Black	2.8	(0.3)
Hispanic	3.1	(0.5)
American Indian	5.2	(1.3)
Asian	2.2	(0.8)
White and Other	2.3	(0.1)
Total Occupied	2.4	(0.2)

### 3.0 The E-Sample

The E-sample housing units were coded as correct enumeration, erroneous enumeration, insufficient information, and unresolved. The estimated percentages in the four categories using data weighted to national totals are in Table 3.1.

A correct enumeration was a census enumeration that was correctly classified as a housing unit according to the census housing definition. An erroneous enumeration was a census enumeration that was in error. The four classifications of an erroneously enumerated housing unit were within block duplicate, nonexistent housing unit, not really a housing unit, and geocoding error. This study estimated the erroneous enumerations to be 2.02 percent of the census.

The classification of insufficient information was defined as an address in the census files without enough information to locate the housing unit for an interview. These housing units had minimal addresses or blank addresses and the maps were missing. These maps were used in the census enumeration and had the location of the housing unit spotted on the map. The minimum amount of information required for HUCS follow-up was a housing unit spotted on the map.

A housing unit was coded unresolved when the field follow-up interview for the Housing Unit Coverage Study was inconclusive. The probability of correct enumeration in the census for the housing units coded unresolved will be imputed in order to calculate dual system estimates. A logistic regression model will be fit to resolved E-sample housing units to predict the probability of correctly enumerated versus erroneously enumerated for unresolved E-sample housing units.

**Table 3.1: E-sample Results Before Imputation**

Correct Enumeration	97.04
Erroneous Enumeration	2.02
Insufficient Information	0.29
Unresolved	0.65
Total	100.00

The estimated percentage erroneous enumeration by type of enumeration area is presented in Table 3.2. Again, only the prelist and update/leave areas were compared because the search area was the same for these two types of enumeration area. The percentage erroneous enumeration for the prelist areas was not significantly different from the update/leave areas.

**Table 3.2: Percentage Erroneous Enumeration by Type of Enumeration Area**

TAR	2.0	(0.2)
Prelist	1.7	(0.2)
Update/Leave	2.7	(1.1)
List/Enumerate	2.4	(0.4)
Total	2.0	(0.2)

The estimated percentage erroneous enumeration is in Table 3.3 for occupancy status in the census. The estimated percentage erroneous enumeration for occupied housing units was significantly different from vacant housing units.

**Table 3.3: Percentage Erroneous Enumeration by Occupancy Status in Census**

Occupied	1.5	(0.1)
Vacant	7.3	(1.0)

The estimated percentage erroneous enumeration is in Table 3.4 for the four census regions. The percentage erroneous enumeration in the Northeast was significantly different from the percentage erroneous enumeration in the Midwest and West regions, but not significantly different from the South

region. The percentage erroneous enumeration in the South, Midwest, and West regions were not significantly different.

**Table 3.4: Percentage Erroneous Enumeration by Census Region**

Northeast	3.0	(0.5)
South	2.2	(0.4)
Midwest	1.2	(0.2)
West	1.6	(0.3)

The estimated percentage erroneous enumeration is presented in Table 3.5 for the five categories of number of units in structure. The category of "Other" included the trailers and any other type of housing unit that was not a structure, such as tents, vans, and boats. The estimated percentage erroneous enumeration in single units was significantly different from the estimated percentage erroneous enumeration for units in 2 to 9 unit structures and for units in 10 to 49 unit structures, but not significantly different from the estimated percentage erroneous enumeration in units in 50 or more unit structures. The percentage erroneous enumeration for "other" was significantly different from the single units, but not significantly different from the three types of multi-unit structures.

The estimate of percentage erroneous enumeration for all multi-unit structures was 2.6 percent with a standard error of 0.5 percent. The estimated percentage erroneous enumeration in multi-units was significantly different from the estimated percentage erroneous enumeration in single units, but not significantly different from the estimated percentage erroneous enumeration in trailers, tents, vans, and boats.

**Table 3.5: Percentage Erroneous Enumeration by Number of Units in Structure**

Single	1.5	(0.1)
2 to 9	2.6	(0.3)
10 to 49	3.0	(1.0)
50 and over	2.4	(1.5)
Other	4.2	(1.0)

The estimated percentage erroneous enumeration is in Table 3.6 for the five race/origin categories for occupied housing units. The race/origin classification was obtained from the first person in the household for the occupied E-sample housing units. The estimated percentage erroneous enumeration for housing units classified as "Black" was significantly different from the housing units classified as "White and Other". The estimated percentage erroneous enumeration for housing units classified as "Black" was not significantly different from the housing units classified as "Hispanic". The estimated percentage erroneous enumeration for housing units classified as "Hispanic" was significantly different from housing units classified as "White and Other".

**Table 3.6: Percentage Erroneous Enumeration by Race/Origin of Person 1**

Black	2.1	(0.3)
Hispanic	2.2	(0.5)
American Indian	1.4	(0.5)
Asian	1.3	(0.4)
White and Other	1.3	(0.1)
Total Occupied	1.5	(0.1)

### 3.1 Reasons for Erroneous Enumerations

The four classifications of an erroneously enumerated housing unit were within block duplicate, nonexistent housing unit, not really a housing unit, and geocoding error. A within block duplicate was a housing unit enumerated twice in the E-sample within the sample block. A nonexistent housing unit was a housing unit enumerated in the block that did not exist in that block or in the search area. The housing

unit may have existed somewhere outside the search area. Some of these may have been geocoding errors that were not located by the field representative because they were incorrectly geocoded a great distance from the sample block. Other census enumerated housing units classified as nonexistent were addresses that were erroneously included in the census files and did not exist.

The enumerations that were not housing units were commercial property, group quarters, or other buildings or structures that should not have been enumerated as housing units. The housing units classified as geocoding errors are housing units enumerated within the sample block that actually existed outside the search area.

The surrounding block duplicates are not discussed in this section and are not reflected in the tables in this section. The percent erroneous enumeration increases when the amount of erroneous enumeration due to duplication in surrounding blocks is added. This section contains a discussion of the four kinds of erroneous enumeration within the E-sample in the sample block.

Table 3.1.1 contains the estimated percent of the census housing units with each of the four reasons for erroneous enumeration in the census. The estimate of the total erroneous enumerations was 2.02 percent of the census.

**Table 3.1.1: Reasons for Erroneous Enumerations**

Within Block Duplicate	0.51
Nonexistent	0.63
Not a Housing Unit	0.43
Geocoding Error	0.45
Erroneous Enumerations	2.02

Table 3.1.2 contains the estimated percentage for each reason for erroneous enumeration as it occurred within each type of enumeration area. The reasons for erroneous enumeration for the four types of enumeration areas can be compared, except for the geocoding error. The definitions for the reasons are the same for all types of enumeration areas, except for the geocoding error.

Housing unit duplication was not significantly different for the four types of enumeration areas. The percentage nonexistent in update/leave areas was 1.29 percent. The TAR areas have geocoding error of 0.68 percent. The percentage of housing units enumerated that HUCS identified as not actually housing units was 0.28 percent in TAR areas and 0.41 percent in prelist areas. The estimated percentage of enumerations that were not housing units was 1.41 percent in list/enumerate areas and 0.85 percent in update/leave areas.

**Table 3.1.2: Reasons for Erroneous Enumerations By Type of Enumeration Area**

Code	TAR	Prelist	U/L	LE
Within Block Duplicate	0.46	0.64	0.40	0.48
Nonexistent	0.58	0.56	1.29	0.35
Not a Housing Unit	0.28	0.41	0.85	1.41
Geocoding Error	0.68	0.13	0.12	0.17

#### 4.0 ACF Evaluation

The data from the Housing Unit Coverage Study was also used to evaluate the quality of the addresses in the Address Control File (ACF). The ACF is a computer file containing the addresses for the housing units enumerated in the 1990 Census. If the existing ACF is used as the base for the address file in the 2000 Census, the quality and completeness of the addresses on the file are important. We need to determine the portions of the file that need to be enhanced before the file is used to mail census questionnaires for the 2000 Census. The census housing units classified as correctly enumerated in the HUCS matching were used to

evaluate the addresses on the ACF. HUCS estimated that 97.0 percent of the addresses in the census were correctly enumerated.

The quality of the census address information in the ACF was recorded while assigning the match codes in the HUCS matching. Table 4.1 presents the correct enumerations from HUCS in the four classifications based on the quality and source of the information used to assign the correct enumeration. This study estimated that 79.7 percent of the correct enumerations were assigned when the HUCS and census addresses matched exactly.

The address was estimated as not the same, but additional information from another source allowed a match to be made between the HUCS address and the census address for 9.6 percent of the correct enumerations. One of the sources for additional information was the address registers in the prelist and update/leave areas. There were cases where the address in the ACF was not keyed correctly from the address register or the address was not keyed at all. The correct address in the address register allowed matching to the HUCS addresses.

In update/leave areas the corrections to the census address from the update operation were not included on the ACF, because there was no operation to key them into the ACF. These corrections were used in the HUCS matching. Also, all addresses for the housing units added from the update operation were designed to be keyed into the ACF. There were cases in the update/leave areas where these added housing units were keyed into the ACF with no address. Using the address in the address register allowed matching to the HUCS address.

There was also additional information in the address register that was not on the ACF such as the location description or the street address in cases where the mailing address was not a house number and street name address or the street address was blank. A keying option in the data entry program could bypass the location description completely, if the keyer incorrectly selected the house number and street name option when the address was a rural address. Also, the ACF limits the number of characters keyed from the address register for location description.

Another source of additional information was the enumerator maps in prelist, update/leave, and list/enumerate areas. The location of the housing unit was spotted on these enumerator maps. When the addresses were not exactly the same, but the location of the housing units on the PES and census maps identified the same housing unit, this information allowed matching the two housing units.

There also were census questionnaires containing address corrections on the label of the questionnaire that were helpful in matching addresses. This usually happened when the respondent corrected the address on the census questionnaire and returned the questionnaire by mail. The ACF was not updated with these address corrections. Also, address corrections obtained during the HUCS follow-up of not matched census housing units were used to match HUCS and census housing units.

The census housing units were classified as matched when the basic address was the same, but the apartment designation was noncontradictory in 4.4 percent of the correct enumerations. Noncontradictory is defined as the HUCS and census apartment designations were different, but the number of units in the building was the same and there was no evidence that suggests these apartment designations do not refer to the same housing unit.

The census housing unit did not match to the PES and was determined to be correctly enumerated during the HUCS field follow-up for 6.3 percent of the correct enumerations. All census housing units not matched to HUCS housing units were followed up in the field to determine correct or erroneous enumeration on April 1, 1990.

Matched exactly to ACF	79.7
Matched using address register, map spotted maps, location description, or other materials	9.6
Multi-unit without matching unit number, but the units are noncontradictory	4.4
Not Matched, but correctly enumerated	6.3
Total Correct Enumerations	100.0

These classifications indicating quality of the address on the ACF for the correct enumerations in the four types of enumeration areas are in Table 4.2. This table gives us an indication of the differences in enumeration areas.

The estimate of the percent of correct enumerations assigned with the help of other materials not on the ACF was 40.8 percent in update/leave areas and 39.1 percent in list/enumerate areas. These are rural areas in which the quality of the addresses existing on the ACF should be updated before the 2000 Census. The reliance upon map spotted maps and other materials may have been because of the incompleteness of the address in the most rural areas of the country. The nature of the enumeration in these areas may have added to the problem, because the census enumerators were not concerned with getting a complete address for mailing the census questionnaires in update/leave areas. These census questionnaires were delivered by census enumerators instead of by the United States Postal Service.

The census processing of the ACF in TAR and prelist areas was designed to incorporate the corrections and updates from pre-census operations into the ACF. The corrections to the address after mailing census questionnaires were not incorporated into the ACF. The HUCS matching operation used these post-census materials to verify the census addresses. The pre-census materials with updates were not verified. As a result, this estimate could be an underestimate in areas where the pre-census corrections were not keyed correctly.

The addresses in the list/enumerate areas were not on the ACF in the census. The matching for HUCS was done with the actual census address registers. The results in the list/enumerate areas are included in Table 4.2 to illustrate the quality of these addresses as compared with the other types of enumeration areas. The list/enumerate addresses were keyed recently and do now exist on the ACF. Any keying errors for the list/enumerate addresses on the ACF are not reflected in this analysis.

In TAR areas 6.7 percent of the correct enumerations in the census were because the unit designation for a multi-unit structure was not the same, but was matched because the units were non-contradictory. This will present problems when the units must be interviewed in the field. If the census questionnaire is not mailed back, a field enumerator visits the housing unit to conduct an interview. If the enumerator can not find the correct apartment to interview, there is a chance of interviewing the wrong housing unit. For example, the mailing list purchased by the Census Bureau had apartments 1 through 10, but the actual apartment designations are 101, 102, 103, 201, 202, 203, 301, 302, 303, and 400. The Census Bureau mailed ten census questionnaires to the ten unit building and nine were returned. Apartment number 6 was not returned in the mail, because the apartment is vacant. The census enumerator visits the building and must decide which unit to visit. If the enumerator picks the wrong unit to interview, the persons may be duplicated.

Table 4.2: Percentage Correct Enumerations by Type of Enumeration Area

Code	TAR	Prelist	UL	LE
Matched exactly to ACF	87.2	78.5	50.9	52.2
Matched using address register, map spotted maps, location description, or other materials	0.4	13.2	40.8	39.1
Multi-unit without matching unit number, but the units are noncontradictory	6.7	1.1	1.0	1.6
Not Matched, but correctly enumerated	5.7	7.2	7.3	7.1

There are sections of the country where the route and box number addresses are being updated or have already been replaced by house number and street name addresses. These rural addresses will be obsolete in the year 2000. These new addresses being added for ease of location by emergency vehicles need to be incorporated into the ACF before the 2000 Census.

## 5.0 Evaluation of Collection Geography

The codes assigned to each E-sample housing unit in the Housing Unit Coverage Study were redefined to reflect the five classifications representing the accuracy of assignment to census geography. The sample housing units were classified as "correct", "incorrect - within the search area", "incorrect - outside the search area", "insufficient information for field interviewing", and "unresolved".

The results of the evaluation of collection geography are in Table 5.1. The collection geography for the housing unit was determined from this study to be correct for 93.92 percent of the housing units and incorrect for 4.54 percent of the housing units. The classification of incorrect had two classifications: "incorrect - within the search area" and "incorrect - outside the search area".

For TAR areas, "incorrect - within the search area" was defined as "the correct geography was one block from the enumerated geography" and "incorrect - outside the search area" was defined as "the correct geography was more than one block from the enumerated geography". For prelist areas and update/leave areas, "incorrect - within the search area" was defined as "the correct geography was one or two blocks from the enumerated geography" and "incorrect - outside the search area" was defined as "the correct geography was more than two blocks from the enumerated geography". For list/enumerate areas, "incorrect - within the search area" was defined as "the correct geography was inside the address register area (ARA)" and "incorrect - outside the search area" was defined as "the correct geography was outside the ARA".

Table 5.1: Evaluation of Collection Geography

Correct	93.92
Incorrect - within search area	4.09
Incorrect - outside search area	0.45
Insufficient Information	0.29
Unresolved	1.25
Total	100.00

The classifications of "incorrect - within the search area" and "incorrect - outside the search area" were combined to reflect the accuracy of coding to the sample block. The estimated percentage of housing units with incorrect census geography is displayed in Table 5.2 by type of enumeration area. Estimated standard errors for the estimates of percentage incorrect geography are in parentheses. The estimated percentage with incorrect geography in both the TAR and prelist areas was significantly different from the estimated percentage with incorrect geography in the list/enumerate

areas. The estimated percentage with incorrect geography in the update/leave and the list/enumerate areas was not significantly different. The estimated percentage with incorrect geography in neither the TAR nor prelist areas was significantly different from the estimated percentage with incorrect geography in the update/leave areas.

**Table 5.2: Percentage Incorrect Geography by Type of Enumeration Area**

TAR	4.7	(0.8)
Prelist	4.9	(1.1)
Update/Leave	3.5	(0.8)
List/Enumerate	2.6	(0.5)
Total	4.5	(0.5)

The estimated percentage of housing units with incorrect geography in the four census regions is presented in Table 5.3. The estimated percentage of housing units with incorrect geography in the South and in the Midwest were significantly different.

**Table 5.3: Percentage Incorrect Geography by Region**

Northeast	3.6	(0.6)
South	6.9	(1.3)
Midwest	2.2	(0.6)
West	4.2	(1.1)

## 6.0 Comparison with 1980 Results

The designs of the 1980 and 1990 coverage studies were different. The 1980 P-sample was a sample of housing units from the Current Population Survey. The published results of the 1980 HUCS state that the 1980 design may not be the optimum one to provide housing coverage estimates and describe an alternate design which is similar to the 1990 design. These 1980 results also describe the limitations of the data. One of these limitations is the estimates of the gross under-enumeration reflect the completeness of the census address registers and not the data tapes from which the final census counts are generated. The 1990 census was more automated. The comparison for the 1990 HUCS was to the actual enumerations in the 1990 census and not to a paper listing that may or may not have been captured in the census where these housing units were listed on paper.

The 1980 and 1990 percentage not matched are presented in Table 6.1. The percent not matched was 3.4 percent for the 1990 HUCS compared to 2.6 percent for 1980. The design for 1990 had a fixed search area that did not allow searching beyond the search area. The search area was not fixed for the 1980 HUCS and errors in searching outside the search area were committed. A balanced search area is required to produce dual system estimates for housing units. No net estimates of coverage could be made from the 1980 HUCS.

**Table 6.1: Comparison of Percentage Not Matched for the 1980 and 1990 HUCS**

	1980 HUCS		1990 HUCS	
Occupied	1.5	(0.1)	2.4	(0.2)
Vacant	12.6	(0.8)	11.4	(1.0)
Total	2.6	(0.1)	3.4	(0.2)

Another limitation of the data for 1980 was that the estimates from the E-sample are over-enumerations and not erroneous enumerations. Over-enumerations referred to multiple enumerations. The sample for the 1980 HUCS was selected from census enumerations of housing units which had over-enumerations of at least one household member. In addition, the units in one portion of the sample were limited to those which had some evidence of possible geographic coding problems. The published results of the 1980 HUCS

stated that the estimates of over-enumeration were an underestimate of the gross multiple enumeration rate, because the sample did not include the vacant housing units or the occupied housing units for which no household member was over-enumerated and for which there was no indication of geographic coding problems. The percentage estimated to be an over-enumeration was 0.9 percent in the 1980 HUCS.

An erroneous enumeration for the 1990 design included all multiple enumerations, duplicates, geographic errors, nonexistent housing units, and nonresidential enumerations regardless of the occupancy status of the unit or the enumeration status of the persons. The estimated percentage erroneously enumerated from the 1990 HUCS was 2.0 percent.

## 7.0 Future Research

Additional research is planned to further investigate the coverage of housing units for the 1990 Decennial Census. We plan to combine the person coverage from the 1990 PES with the housing coverage from the Housing Unit Coverage Study. This will allow us to investigate the person coverage in the P-sample and the E-sample housing units for the Housing Unit Coverage Study. For example, we will be able to investigate the duplicate persons and housing units. We will be able to estimate the number of persons duplicated because of duplicated housing units and the number of duplicated persons in correctly enumerated housing units. We will also be able to investigate the correctly enumerated persons in not matched housing units and the number of not matched persons in not matched housing units.

We also plan further research for the evaluation of collection geography. We plan to study the different methods of address compilation for the different types of enumeration areas.

These data will help us to learn more about the 1990 census of housing units. It is important for the Bureau to investigate areas where the census methods were successful and areas where the methods need improvement in preparation for the next decennial census. The Census Bureau may not have much control over within household persons missed or within household persons erroneously enumerated, but we do have control over including all housing units once and only once in the census files.

## 8.0 References

"The Coverage of Housing in the 1980 Census", Evaluation and Research Reports, 1980 Census of Population and Housing, PHC80-E1, July 1985.

Hogan, Howard (1990), "The 1990 Post-Enumeration Survey: An Overview", Proceedings of the Section on Survey Methods Research of the American Statistical Association, pp 518-523.

Hogan, Howard (1991), "The 1990 Post-Enumeration Survey: Operations and Results", Proceedings of the Section on Survey Methods Research of the American Statistical Association.

Hogan, Howard and J. Gregory Robinson (1992), "Measuring and Correcting for Census Coverage Errors", Presented at the 14th Population Census Conference, May 26-29, Tokyo, Japan.

Woltman, Henry, Alberti, Nick, and Moriarity, Chris, "Sample Design for the 1990 Census Post Enumeration Survey", 1990 Decennial Census Preliminary Research and Evaluation Memorandum No. 16, September 19, 1989.

\*This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.