

# THE 1988 POST ENUMERATION SURVEY METHODS AND PRELIMINARY RESULTS

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

**KEY WORDS:** Post Enumeration Survey,  
Undercount, Overcount, Adjustment

## 1.0 Introduction

The Dress Rehearsal Post Enumeration Survey (PES) served two purposes. It was both a Dress Rehearsal of the 1990 PES and an evaluation of the Dress Rehearsal Census. The 1988 PES consisted of two samples. The first was a population or P-sample, which consisted of all people living in the PES sample blocks at the time of the PES interview. The other sample was the enumeration or E-sample, which consisted of all enumerations assigned to the PES sample blocks by the census process.

In general, the 1988 Dress Rehearsal PES followed the steps laid out in earlier tests. The addresses for the housing units in the sample blocks were listed and large blocks were subsampled. The PES interview collected names, characteristics, and census day addresses for all residents in sample households at the time of the PES interview. Also, names and characteristics were obtained for persons moving from the sample address between census day (March 20, 1988) and the PES interview. The Dress Rehearsal PES and census data were computer matched. Following the computer matching, the matching clerks performed within block matching, duplicate searching, and matching in the surrounding blocks. Reported alternate addresses were coded to census geography (i.e., geocoded) and persons were searched for at their census day address.

Selected cases were sent to the field for a follow-up interview and the results of the follow-up interview were recorded on the match forms. When new census day addresses were obtained for PES persons during the follow-up interview, these persons were searched for at the new census day address. The final result was to classify each P-sample person as counted or missed in the census counts. Each E-sample person, in turn, was classified as correctly or erroneously enumerated in the census.

## 2.0 Design

### 2.1 Sample Design

The 1988 PES was designed to be a block sample with an overlapping P-sample and E-sample. The P-sample measures undercoverage and the E-sample measures overcoverage. Persons in the P-sample were compared to persons in the E-sample. All persons found in both the P-sample and the E-sample were classified as matched and correctly enumerated in the census without error.

Persons interviewed in the P-sample and not matched to the census were considered missed in the census. Persons in the E-sample who were not matched to the P-sample were reinterviewed to determine if they should have been classified as correctly or erroneously enumerated in the census. Overcoverage resulted when a person was erroneously enumerated in the census, where an erroneous enumeration was any person or household counted

incorrectly. These errors arose from duplicate enumerations, persons enumerated in the wrong household according to census residence rules, and housing units enumerated in the wrong geographic area. The difference derived from subtracting undercoverage and overcoverage has been denoted as the net coverage error of the census.

### 2.2 The Search Area

Persons in the P-sample were also compared to non E-sample census enumerations in the search area. P-sample persons matching these non E-sample census enumerations were also classified as correctly enumerated in the census.

The search area was the distance in geographic area for searching in the census for a match to the P-sample persons and for duplicates and erroneous enumerations in the E-sample. This distance for searching was determined by the PES design before matching began. The size of this distance was theoretically unimportant in terms of expected value of the net undercount. The only requirement was to make the P-sample and E-sample search area consistent. The search area was the sample block or block cluster and the census blocks surrounding the sample block or block cluster. The size of the search area will determine cost and variance. As the size of the search area increases the cost increases and the variance of the estimate decreases.

The surrounding blocks in the 1988 PES for tape address register (TAR) areas consisted of one ring of blocks around the sample block or block cluster. The surrounding blocks for prelist and update/leave areas consisted of two rings of blocks around the sample block or block cluster. The first ring of blocks in the search area was all blocks touching the sample block or block cluster at one or more points. This included the corner blocks, which were blocks touching at only one point. The second ring of blocks, likewise, was all blocks touching the first ring of blocks at one or more points.

In 1988 the list/enumerate (L/E) areas were sampled by address register area (ARA). A sample of blocks was clerically selected from the ARA for the PES sample. The sample blocks in the ARA were given a cluster number and were matched by cluster as in all other types of enumeration areas. The surrounding blocks for L/E areas were those blocks in the selected ARA not sampled for interviewing in the PES. The search area for L/E areas was the entire ARA.

### 2.3 Unique Address Matching

A single census day address was assigned to all PES persons. This address was assigned by applying census residence rules to the information reported by the person at the time of the PES interview. The P-sample person was searched for only at this census day address and in the corresponding search area. The PES interview determined where a P-sample person should have been counted. A person was declared a

match if enumerated at that address. For the majority of cases, the correct address was the only address obtained. All movers were searched for in the census at their reported census day address.

Uncertain cases occurred when alternate addresses were obtained in the probes for alternate addresses during the PES interview. The types of alternate addresses were at a college or university, on a military base or ship, at a second home, with another relative, in the hospital, or somewhere else for any reason. Census residence rules were applied to the types of addresses and the dates they stayed at the addresses. The matching operation searched for the PES sample persons at the addresses where the persons should have been counted according to census residence rules.

E-sample cases were coded as correctly enumerated only if the person should have been counted in the sample block or in the search area. Persons counted in the sample who should have been counted elsewhere were called erroneous enumerations.

#### **2.4 Insufficient Information For Matching**

We have found that certain classes of PES interviews could not be matched in an unbiased way. These were cases with insufficient information to declare a person as not enumerated if not matched. Similarly, certain census enumerations lacked sufficient information to determine a match even if the same person was included in the PES. No matching was attempted for these cases with insufficient information for matching. Instead they were treated as missing data.

A PES or census person had insufficient information for matching when the name was incomplete or there were not enough data items to be declared enumerated in the census or interviewed in the PES. A complete name was defined as consisting of both first name and surname.

#### **3.0 Address Listing**

Independent listings of addresses for all structures in the sample blocks were compiled in the PES Address Listing Books in February 1988. Interviewers from the regional offices who conduct the current survey interviews were employed to conduct the 1988 PES address listing to assure independence between the PES and the census.

#### **4.0 Interviewing**

An interview was conducted at each sample household using the PES interview form. The names and characteristics were recorded for all persons living at the sample addresses at the time of the PES interview in July and August 1988. All persons were asked how long they had lived at the address. If they lived at the address less than one year, they were asked when they moved to the address.

If they moved there since census day, they were asked where they lived on census day. Persons who moved to the address since census day were considered in-movers. Their census day addresses were recorded on the PES interview form. If they lived at the addresses longer than one year or if they moved to the addresses before census day, they were considered non-movers. All persons were also asked if there were

any other places where they might have been counted in the census to determine unique census day addresses.

The first three weeks of interviewing were restricted to interviews with a household member. During the fourth week, interviews with nonhousehold members were accepted. Last resort interviews were accepted during the final week of interviewing. Great effort was made to obtain a complete interview with a household member. Proxy information from a neighbor or apartment manager had not been shown to be reliable in past tests. Persons not residing in the sample household do not have the best knowledge of the census day residence of a sample person. The estimate of the number of persons missed in the census would not be reliable when the incorrect census day addresses were obtained for sample persons.

### **5.0 Matching**

#### **5.1 Computer Matching**

The primary goal of the computer matching system was to obtain as many non-erroneous matches as possible while substantially reducing the workload of clerical matching. This goal was highly dependent upon the quality of the computer data obtained through interviewing, transcribing, and keying.

Computer matching was only performed for those pairs agreeing on logical blocking characteristics such as census block number and a numerical code for the surname. This restriction reduced computation by many orders of magnitude. Most pairs that were ultimately matched agreed on the code of the surname. Matches and possible matches were delineated by way of an algorithm that used individual identifying characteristics such as first name, age, and house number.

#### **5.2 Clerical Matching**

##### **5.2.1 Within Block Matching**

All persons reported in the PES interview in the sample block were computer matched to the persons enumerated in the census in the sample block. This included the non-movers, the in-movers, the out-movers, persons born since census day, and persons with no mover/non-mover status.

The matching for out-movers in the sample block was not part of the PES estimation. They were included in the matching to reduce the amount of E-sample follow-up. All out-movers listed in the PES interview should have matched, because they were reported as living at the sample address on census day, but moved before the PES interview. Out-movers matched to census persons within the sample block confirmed they were correctly enumerated in the census. No follow-up was required for the out-movers who matched.

The in-movers were included in the within block matching, even though they reported living at another address on census day. A match to the census for the in-mover confirmed the in-mover's erroneous enumeration in the census in the sample block. E-sample follow-up for the PES in-mover matching to the census in the sample block was not necessary, since the in-movers told us in July that they did not live at the sample address on census day.

Persons with undetermined mover/nonmover status were sent to follow-up to collect more information about their census day residence. The status of these persons was not based on whether they matched to the census or not. We must determine their correct census addresses independent of their match status to the census.

The sample addresses with "conflicting" households have been referred to as the Smith/Jones or Emerson/Peterson cases. These were households where the PES and census addresses are the same, but the persons reported in the PES and in the census were not the same. These "conflicting" households were targeted for follow-up to determine which household actually lived at the sample address on census day.

#### **5.2.2 The Surrounding Block Match**

The extent of PES search defined the blocks included in the surrounding block match. Any PES sample person enumerated in the census within the PES search area was defined as matched and correctly enumerated in the census. It follows that the surrounding blocks were searched only for P-sample nonmovers who were not matched within the sample block.

The computer match in the surrounding blocks and in the sample block were done at the same time. Matches and possible matches in surrounding blocks were printed on the match form. The computer matching was clerically reviewed, followed by a clerical search in the surrounding blocks for matching persons. Additional matches and possible matches discovered by the clerks were transcribed to the match form.

#### **5.2.3 Duplicate Search**

Census duplication occurred within the sample block and between blocks. The search for census duplicates was conducted within the search area. This census search area was identical to the P-sample search area to maintain the balance between the P-sample and the E-sample. The clerks reviewed the census listings for the sample block and all blocks in the surrounding blocks. The listings in name order help locate duplication with an E-sample person. The listings in household order by address help in locating address duplication between and within blocks.

#### **5.2.4 The Alternate Address Match**

The movers who reported an alternate census day address on the PES interview form were searched for in the census. This clerical match included geocoding, address searching, and person searching. The computer matcher could not be used because the census names were not keyed for the entire test site.

The alternate census day addresses were geocoded using the block header record, the online computer file of census addresses, and computer printed listings of addresses. Maps were also needed for addresses that could not be geocoded from the census address file and for rural addresses. There was re-geocoding for addresses not geocoded correctly and geocoding for new addresses obtained during the follow-up interview.

The census address file for the specific census geography was searched for the alternate census day addresses. If the address was matched in the census address file, the ID number identified the census questionnaire for person searching. If the address was not matched, but fell within the house number range for the street names in the block or was a multi-unit within the block, a range of ID numbers identified the census questionnaires for person searching. In both cases the geography was confirmed and the address should have been enumerated in the block. The PES names were searched for on copies of the census questionnaires or range of questionnaires. The copies were generated from microfilm readers, since the census questionnaires were not sorted to census geography.

The person matching rules were applied in the alternate address matching. If a matching person was not found on the census questionnaire, the census questionnaires for the entire block were searched. The next step was to search the surrounding blocks. The type of enumeration area for the block was identified, along with the corresponding surrounding blocks. If the search was unsuccessful, the PES person at the alternate address was coded not matched.

If the geography was not confirmed, the case was sent to be re-geocoded and the correct geography was determined. Additional information was collected during the PES interview to aid in confirming the census geography, such as cross streets, landmarks, and neighbor's names. Accurate census geography was required in order to classify a person as missed in the census. A person could not be coded as not matched until the geography was confirmed. If the geography could not be confirmed as correct, the person was sent for a follow-up interview.

#### **5.2.5 Matching Late Census Data**

Some census data were obtained after the computer matching began for a district office (DO). There were several reasons for this additional data. The search/match operation did not end until after the PES matching started. Other census questionnaires were not data captured in time to be included. Other census operations such as post census local review added, deleted, or moved housing units to the census after PES matching began.

The matching was designed to start when the DO was approximately 98% to 99% complete. This also allowed us to start computer matching in time to meet the schedule for the PES. The late data were included, however, for the PES estimates to be accurate. The late data were clerically added to the matching before the follow-up forms were prepared. All PES matching operations were performed for the added persons. These new census persons could be E-sample persons, non E-sample persons in a sample block, or non E-sample persons in a surrounding block.

When a new address was added to the census, the PES persons could have been matched. If they were not matched, the address could have matched, requiring a change in the type of nonmatch code. One person could have been matched and the other household members could have become partial household nonmatches. One or more persons could have been added to an existing census household.

Duplicate search was performed on all census persons added to the search area.

### 5.3 Computer and Clerical Matching Results

Tables 1, 2, and 3 display the results of the computer and clerical matching before the follow-up operations for the three district offices in dress rehearsal. The computer match code was the code assigned after computer matching all nonmovers. The final before follow-up match code was the match

code assigned after all phases of the clerical matching. These phases consist of within block matching, duplicate searching, and surrounding block matching by the clerical matching group and the special matching group. The results of the quality assurance adjudication of differences in the production and quality control matching were also included, as were the adds and deletes from processing the census data obtained after computer matching began.

**Table 1: St. Louis P-sample  
Nonmovers Before Follow-up**

Before Follow-up Match Code	Computer Match Code					Total	%
	Match	Nonmatch	Possible Match	Incomplete			
Match	8,763	921	591	12	10,287	85.8	
Nonmatch	11	1,345	176	1	1,533	12.8	
Possible	3	30	43	0	76	0.6	
Unresolved	8	69	14	8	99	0.8	
Total	8,785	2,365	824	21	11,995		
%	73.2	19.7	6.9	0.2			

**Table 2: Columbia P-sample  
Nonmovers Before Follow-up**

Before Follow-up Match Code	Computer Match Code					Total	%
	Match	Nonmatch	Possible Match	Incomplete			
Match	5,960	505	648	3	7,116	89.4	
Nonmatch	2	438	29	0	469	6.2	
Possible	0	5	8	0	13	0.2	
Unresolved	2	17	1	4	24	0.3	
Total	5,964	965	686	7	7,622		
%	78.3	12.7	9.0	0.1			

**Table 3: Washington P-sample  
Nonmovers Before Follow-up**

Before Follow-up Match Code	Computer Match Code					Total	%
	Match	Nonmatch	Possible Match	Incomplete			
Match	1,711	157	86	5	1,959	89.4	
Nonmatch	1	215	1	0	217	9.9	
Possible	0	8	2	0	10	0.5	
Unresolved	0	6	0	0	6	0.3	
Total	1,712	386	89	5	2,192		
%	78.1	17.6	4.1	0.2			

The computer matcher coded 73.2 percent of the nonmovers as matches in St. Louis, 78.3 percent in Columbia, and 78.1 percent in Washington. These differences should reflect the expected results in 1990 by type of district office. St. Louis is a type 1 district office, Columbia is a type 2 district office, and Washington is a type 3 district office.

Type 1 district offices consisted exclusively of TAR areas, which were urban areas where the Census Bureau purchases mailing lists. Large cities and inter-city areas in type 1 district offices have typically been hard to enumerate. Type 2 district offices were composed of TAR, prelist, and update/leave areas. Type 2 district offices were mixed with smaller cities and rural areas. Type 3 district offices were composed of TAR, prelist, and list/enumerate areas. Type 3 district offices were mostly rural including small towns near the rural areas. Assuming the three district offices in dress rehearsal were representative of the three types of district offices for 1990, we should expect a computer

match rate of approximately 78 percent in type 2 and 3 district offices and 73 percent in the harder to enumerate type 1 district offices.

After the clerical matching, 99.75 percent of the matches assigned by the computer matcher were classified as matches in St. Louis, 99.93 percent in Columbia, and 99.94 percent in Washington. In St. Louis 0.13 percent of the matches assigned by the computer matcher were classified as nonmatches, 0.03 percent in Columbia, and 0.06 percent in Washington.

Subtracting the percent matched by the computer and the percent matched after the clerical review will not result in the percent matched by the clerical matchers. Some matches assigned by the computer were not allowed to remain as matches. Some of the matches were not valid and the real match was discovered during the clerical matching, which is not reflected in these tables.

These before follow-up data cannot be used to calculate nonmatch rates, because the results of the

follow-up interview are not included. Further, the movers are not represented in these data. Any differences in these data and any subsequent data result from the follow-up interview and after follow-up matching.

#### 5.4 Follow-up

The follow-up for the Dress Rehearsal PES was designed to obtain additional information to complete the matching operations. Its goals were to reduce the size of the nonresponse, unresolved, unreliable match, and unreliable nonmatch categories. In general we sent to follow-up only PES cases with questionable, insufficient, or missing information. We did not send PES cases to follow-up simply because they did not match to the census enumerations. All census nonmatches were followed up to determine correct or erroneous enumeration in the census.

#### 5.5 After Follow-up Matching and Coding

The results of the follow-up interview were recorded on the match forms and entered into the computer system. The final result was to classify each P-sample person as counted or missed in the census counts. Each E-sample person was classified as correctly or erroneously enumerated in the census.

#### 5.6 Preliminary Results

The results and tabulations in this paper are preliminary. The numbers reflect raw data without weighting and imputations. They are presented and discussed as part of the process of understanding the results, both substantive and methodological, of the Dress Rehearsal PES.

##### 5.6.1 P-sample by Mover/Nonmover

Tables 4, 5, and 6 contain the percent of P-sample persons in housing units in Columbia, St. Louis, and Washington for movers and nonmovers by final match code.

We were only able to match movers within the test site for dress rehearsal, because we did not have a census of the entire country. Columbia was chosen as a part of the dress rehearsal to include an area that would contain colleges and seasonal housing units. St. Louis was chosen to test a hard-to-enumerate inner city area. The test site in Washington was selected to test a rural area for 1990.

Comparing the three test sites, the total percent nonmatch in St. Louis was about twice the percent nonmatch in Columbia. The total percent nonmatch for Columbia was 6.2 percent, for St. Louis the percent nonmatch was 13.0 percent, and for Washington the percent nonmatch was 10.1 percent.

**Table 4: P-Sample Persons in Housing Units in Columbia in Percent**

	<u>Mover</u>	<u>NonMover</u>	<u>Total</u>
Match	56.5	93.6	91.1
Nonmatch	8.5	6.1	6.2
Unresolved	<u>35.0</u>	<u>0.3</u>	<u>2.7</u>
Total	6.9	93.1	100.0

**Table 5: P-Sample Persons in Housing Units in St. Louis in Percent**

	<u>Mover</u>	<u>NonMover</u>	<u>Total</u>
Match	52.9	86.8	85.3
Nonmatch	26.5	12.4	13.0
Unresolved	<u>20.6</u>	<u>0.9</u>	<u>1.7</u>
Total	4.3	95.7	100.0

**Table 6: P-Sample Persons in Housing Units in Washington in Percent**

	<u>Mover</u>	<u>NonMover</u>	<u>Total</u>
Match	67.2	89.8	88.0
Nonmatch	13.3	9.8	10.1
Unresolved	<u>19.4</u>	<u>0.4</u>	<u>1.9</u>
Total	7.7	92.3	100.0

The nonmover percent unresolved was less than one percent in all three sites. The percent unresolved for movers in Columbia was almost 35 percent, in St. Louis about 21 percent, and almost 20 percent in Washington. These unresolved cases for movers were composed of problems in geocoding the mover addresses received within the test site and of cases geocoded, but the other information like names of the neighbors and cross streets was not confirmed during the matching.

The persons not matched represent cases for which we were confident that the mover did not get counted at their unique census address according to census residence rules. This does not mean that a person was missed in the census. It means the movers were not enumerated where they should have been counted according to census residence rules.

##### 5.6.2 P-sample Nonmovers by Type of Nonmatch

The nonmatches for nonmovers were coded by type of nonmatch. The code "L" indicated the nonmover was matched to a person on a census questionnaire that was rejected during census processing. This situation usually occurred when there were two census questionnaires data captured with the same census identification number. One was enumerator filled and one was a late mail return. Only one of the duplicate census questionnaires could have been accepted as the census enumeration for the household. On some occasions the PES persons matched the persons on the census questionnaire that was not accepted. Even though the PES and census persons were matched, they were considered not matched, since the census person was not considered enumerated in the census.

Persons can also be coded "L" within a household from an accepted census questionnaire. This occurs when the matched person was deleted from the census questionnaire by filling the cancellation circles. All PES persons coded "L" are considered not matched to the census.

The code "N1" indicated a partial household nonmatch. There were other persons in the household who were matched or possibly matched to the census. The code "N2" indicated a whole household nonmatch where the address was matched to the census. There existed a census questionnaire

for the address, but other persons or no one was enumerated at the address. The code "N3" indicated a whole household nonmatch where the basic address was enumerated in the census, but the PES apartment or unit was not enumerated in the census. These occur in multi-unit buildings or trailer parks. No census questionnaire was enumerated in the census for this unit. The code "N4" indicated a whole household nonmatch where the PES address was not found in the census. This was a housing unit nonmatch with all persons not matched in the census. The percent not matched by type of nonmatch is in Table 7 for the three sites.

**Table 7: Type of Nonmatch for Nonmovers in Percent**

Type of Nonmatch	Columbia	St. Louis	Washington
L	4.6	3.9	1.4
N1	30.6	39.4	30.3
N2	11.5	40.4	13.3
N3	6.5	6.9	10.9
N4	46.8	9.3	44.1

The percent of the nonmatches for nonmovers who were recorded on a census questionnaire, but later removed from the census counts in Columbia, St. Louis, and Washington were 4.6, 3.9, and 1.4 percent, respectively. The gross percent nonmatch could be reduced by these percent for each site, if these rejected persons were actually enumerated. In dress rehearsal the enumerator filled census questionnaire was kept over the late mail return. This has been changed for 1990. The mail return questionnaires have been proven to be more reliable than the enumerator filled questionnaires.

The housing unit was not found in the census for over half of the persons not matched to the census in Columbia and in Washington (N3+N4). For St. Louis the same percentage was 16.2 percent. The census mailing lists in St. Louis were purchased from vendors. In Columbia and Washington the mailing lists were purchased for the TAR areas and compiled by the Census Bureau for the remainder of the areas.

The rural areas in Columbia and Washington had a high percent nonmatch with no matching census address. If the address matching was done correctly in the more rural areas, these nonmatches represent a high rate of housing unit misses in the census in rural areas. This could also have been person nonmatches where the address matching was not performed accurately. When the results from the Housing Unit Coverage Study (HUCS) are available, we will be better able to access the quality of coding the address matching for person nonmatches.

**5.6.3 E-sample by Strata**

The E-sample results are in Tables 8, 9, and 10 for the three sites by strata. The percent erroneous enumeration was 3.3 percent for Columbia, 4.4 percent for St. Louis, and 3.0 percent for Washington. The duplication between sample persons and non E-sample persons in the search area was not represented in these tables.

**Table 8: E-sample in Housing Units in Columbia by Strata in Percent**

Enumeration Status	TAR	Prelist	U/L	Total
Correct	93.9	95.7	96.1	95.4
Erroneous	3.3	3.3	3.4	3.3
Unresolved	2.8	1.0	0.5	1.2

**Table 9: E-sample in Housing Units in St. Louis by Strata in Percent**

Enumeration Status	Black Renter	Black Owner	NonBlack	Total
Correct	90.0	92.7	93.7	91.9
Erroneous	5.7	3.7	3.1	4.4
Unresolved	4.2	3.6	3.2	3.8

**Table 10: E-sample in Housing Units in Washington by Strata in Percent**

Enumeration Status	TAR	Prelist	L/E	Total
Correct	96.1	94.8	94.4	94.7
Erroneous	2.4	1.9	3.7	3.0
Unresolved	1.5	3.3	2.0	2.3

In Columbia the percent erroneous enumeration was the same for all three types of enumeration areas. The percent unresolved was largest for TAR areas. In St. Louis the Black-renter stratum had the highest percent erroneous enumeration and the percent unresolved was largest for the Black-renter stratum. In Washington the list/enumerate areas had the largest percent erroneous enumeration and the percent unresolved was highest for prelist areas. St. Louis has the highest percent unresolved of the three sites.

The percentage of erroneous enumerations by types of enumeration are in tables 11, 12, and 13 for Columbia, St. Louis, and Washington. The types of enumeration referred to mail return or enumerator filled census questionnaires. There was no mail return for the list/enumerate areas in Washington. The census questionnaires filled by census enumerators had a higher percent erroneous enumeration than census questionnaires completed by the household and returned to the Census Bureau by mail for all strata in all three sites.

The percent erroneous enumeration was the number of erroneous enumerations divided by the sum of the correct and erroneous enumerations. This percent was the erroneous enumerations as a percent of the resolved census enumerations. The unresolved cases were assumed in this percent to have the same percent erroneous enumeration as the resolved cases.

**Table 11: Percent Erroneous Enumeration By Type of Enumeration In Columbia**

Type of Enumeration Area	Mail Return	Enumerator Filled	Total
TAR	2.8	4.4	3.4
Prelist	2.9	4.3	3.3
Update/Leave	2.6	5.6	3.4
Total	2.8	4.7	3.4

**Table 12: Percent Erroneous Enumeration  
By Type of Enumeration In St. Louis**

Type of Enumeration Area	Mail Return	Enumerator Filled	Total
Black Owner	2.3	5.8	3.9
Black Owner	2.8	8.8	6.0
NonBlack	2.2	5.1	3.2
Total	2.4	7.1	4.5

**Table 13: Percent Erroneous Enumeration  
By Type of Enumeration In Washington**

Type of Enumeration Area	Mail Return	Enumerator Filled	Total
TAR	0.6	4.5	2.5
Prelist	0.8	3.8	2.0
List/Enumerate	NA	3.7	3.7
Total	0.7	3.8	3.1

### 5.7 Remaining Steps

The missing data for the PES was handled in a variety of ways. Noninterviews and whole household duplicates in the P-sample were given a weighting adjustment within block and type of structure (single unit attached, single unit not attached, multiunit, special place, and other). For missing data characteristics (race, tenure, etc.) a hot-deck procedure was used. For age and sex, a more involved imputation procedure used household size and relationships. For missing enumeration status, separate logistic regression models were used for the P-sample and the E-sample.

The dual system estimates of population size were computed separately within post-strata. This method increased homogeneity and reduced correlation bias. Post stratification was based on race, housing tenure, type of enumeration, age, and sex. For St. Louis we used two race groups (white and all other race/ethnic groups) and two tenure groups (owner and nonowner). In Columbia the groupings were white TAR, white non-TAR, and nonwhite (TAR and non-TAR). For Washington two census procedures areas were used (list/enumerate and non-list/enumerate). Within each of these geographic areas, cells based on age (0-9, 10-19, 20-29, 30-44, 45-64, and 65+), and by sex were formed. There were 108 post-strata in all.

The same undercount rates were applied to all people within the post-strata. These assumed rates are called factors. The factors are smoothed using a regression model. This step narrowed the confidence interval due to sampling uncertainty. Estimates of the census undercount were also based on both the demographic analysis estimates from 1980 and the PES. The PES gave the geographic distribution and the characteristics of the missed people. The national sex ratios from demographic analysis were used with the female PES estimates to improve the estimates for males, where the PES was often in error. From both these sources a combined estimate was derived. Applying the factors to the census count for any area in the test sites produced an estimate of the

population and by subtraction an estimate of the undercount. This method of estimation is being evaluated.

The PES results are being evaluated by two matching studies and additional follow-up interviews. The results of the evaluations will be presented as the analysis is completed.

### 5.8 Conclusion

We are in the process of completing the Dress Rehearsal PES. We will continue to study the results and make changes where improvements are indicated within each phase. Streamlining and fine tuning each operation is in progress for the 1990 PES.

The interaction and communication between seven processing offices could not be tested in dress rehearsal. The mover matching requires transmitting materials between processing offices, because we must use the maps and census questionnaires in another processing office when someone moves a large distance between census day and the time of the PES interview.

The timing of each operation requires review in order to meet the new completion date required by the recent court decision concerning adjustment. The timing was not tested during dress rehearsal.

The follow-up interview attempted to identify nonexistent or fictitious persons reported in the PES and in the census. Fictitious persons are difficult to identify. We require contact with three qualified respondents who never heard of the person. They must be interviewed correctly by the follow-up interviewer and processed properly by the matching clerks. We have had difficulty in the tests in getting the clerks to code as someone as a fictitious person. Work will continue for 1990.

The mover matching also requires attention before 1990. Some of the reported alternate census day addresses for movers were not as complete as we would like. These addresses were difficult to code to census geography when the address was incomplete. We also asked for the names of neighbors and cross streets of the mover address. When the names of neighbors and cross streets were confirmed for the mover address, we were confident that the alternate address has been geocoded correctly. Without this confirmation, we were skeptical of the census geography assigned for mover searching. We are introducing new procedures for 1990 to indicate the confirmation of census geography.

Still, we are pleased with the overall design as tested. What is needed is proper implementation in 1990.

---

\*This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau. A detailed version of the paper can be obtained by writing to Dr. Danny R. Childers, Undercount Research Staff, Room 3000, Bldg. 4, Bureau of the Census, Washington, DC 20233.