

MATCHING MOVERS IN A POST-ENUMERATION SURVEY: RESULTS FROM 1988 AND PLANS FOR 1990*

Irwin Anolik, Bureau of the Census

ABSTRACT

Studies of other censuses have confirmed that persons moving at a time close to Census Day are at greater risk of being omitted from the census or of being enumerated at a subsequent address rather than at their correct Census Day address. Movers are defined in the Post-Enumeration Survey (PES) as persons who report a Census Day address that is different from their PES address. This paper discusses the results of matching movers in the 1988 Dress Rehearsal PES. Before a mover can be matched at his or her Census Day address, the address must be assigned census geography or "geocoded". If a mover cannot be matched at the geocoded address, then census questionnaires are searched within a pre-defined area around that address. This paper also discusses how current plans for matching movers in the 1990 PES were influenced by experiences in 1988.

1. INTRODUCTION

The 1988 Dress Rehearsal PES was a coverage measurement survey conducted after the 1988 Dress Rehearsal Census was carried out in St. Louis City, East Central Missouri and Eastern Washington State. It was designed to be a Dress Rehearsal of the 1990 PES. Another objective was to evaluate person coverage of the Dress Rehearsal Census for such demographic characteristics as age, race, sex, owner-renter status and method of questionnaire delivery. Undercount estimates for these characteristics are found in the paper by Childers and Hogan (1990).

The basic design of the 1988 PES involved an attempt to match persons sampled in the PES (P-sample) to their Census Day address where they should have been enumerated by the Census. An important proportion of the P-sample are "movers" who report a Census Day address that is different from their PES sample address. In other words, they moved between census day and the time of the PES interview.

This paper discusses the results of matching these movers as part of the 1988 Dress Rehearsal PES. We include a discussion of our attempts to assign census geography to movers' Census Day addresses. This so-called "geocoding" operation is an essential prerequisite to matching movers.

This paper also discusses plans for matching movers in the 1990 PES.

Before discussing the above topics, we will summarize two design decisions made for the 1988 Dress Rehearsal PES that directly affected how mover matching was conducted.

2. DESIGN DECISIONS AFFECTING MOVER MATCHING

1988 PES mover matching was affected by two design decisions. One decision was to use a so-called "PES B" procedure to determine match/nonmatch status. In this procedure, the PES interviewer lists all the persons living or staying in the housing unit at the time of the PES. The PES information for nonmovers is matched with the census. In-movers (persons who moved into the sample block between Census Day and the PES interview) are asked where they lived on Census Day. Their Census Day address is searched in attempting to match in-movers to the census. If their Census Day address is outside the test site, then the person is coded as being out-of-scope.

The major alternative to the PES B approach is called "PES A." The PES A procedure reconstructs the households as they existed at the time of the census. It attempts to obtain names and basic characteristics of persons who moved out (out-movers) between Census Day and the time of the PES interview. In either case the PES information is then matched with the census data. The difference between PES A and PES B involves people who move between Census Day and the time of the PES interview.

The PES B procedure was chosen for the 1988 PES because it reduces the need to get information from neighbors or from other non-household members as to who was liv-

ing in the housing unit at the time of the census. However, it requires that in-movers give complete and accurate information on where they were living at the time of the census. This information is used in searching for the persons in the census listings at these former locations.

The second design decision affecting matching involved determining the extent of search. We decided to use an approach referred to as "correct address matching" which searches the census files where the person should have been enumerated in the census. The P-sample person is coded as a match when (s)he is enumerated at the correct census day address, as determined by census residency rules. Otherwise, a non-match is assigned. Depending on the type of enumeration area a search area was defined around each address. In the Tape Address Register¹ (TAR) areas, this search area was the block containing the address and one ring of blocks surrounding that block. In the prelist² areas, this search area was the block containing the address and two rings of blocks surrounding that block. In List/Enumerate³ (L/E) areas, the search area was the block containing the address and the remainder of the ARA. Note that for movers, matching searches have to be made outside PES sample areas.

For the 1990 PES, the design decisions discussed above have not changed.

3. GEOCODING AND MATCHING

Matching is especially difficult for PES persons who lived elsewhere on census day, i.e., movers. Before a mover could be matched at his or her Census Day address, the address had to be coded to census geography (geocoded). Census questionnaire data would then be searched at the mover's Census Day address. If any movers remained to be matched at such an address, census questionnaire data would be searched within a pre-defined area around that address.

Establishing the correct census geography in which to search for a mover was often problematic. The difficulties were more acute in rural areas due to the nature of many rural type addresses and to the limited resources available for assigning census geography to a given address.

In some cases, PES respondents did not know or refused to give Census Day ad-

resses for movers within the PES household. In other cases, a reported Census Day address proved insufficient to either establish the correct census geography or to limit the area of search to a manageable size.

Many respondents did report Census Day addresses considered sufficiently complete to search. However, the initial results of mover matching showed a high non-match rate suggesting the possibility that a substantial fraction of movers reported Census Day addresses that were either incorrect or mistakenly coded to the wrong ARA and block. Hogan and Wolter (1988) point out that inaccurate reporting of census day addresses is one of the main sources of error in a PES. The consequences of such errors, as pointed out in Fay et al (1988), would be to overestimate systematically the nonmatch rate to the census for movers. Thus clerks reviewed movers in the 1988 PES who were either nonmatched, unresolved or out-of-scope. Unresolved cases include movers whose Census Day addresses could not be geocoded. Out-of-scope refers to movers who moved from addresses outside the test site. For 1990, out-of-scope refers only to movers who lived outside the U.S. on census day.

For this review, more resources were made available to the geocoding clerks who, this time, were instructed to use these resources to attempt to find cross streets, roads, highways or other landmarks closest to the movers reported Census Day address for each of these addresses that could be geocoded. In addition, matching clerks were instructed to examine the census questionnaire data to find and thus confirm reported neighbors and census day roster for movers that could be geocoded, but could not be matched to the census. If none of the above information could be found, or if none was available, the movers remaining unmatched would be considered as unresolved since we could not confirm that we were searching for them in the right place.

We incorporated any changes resulting from this review into a file of "final" match codes. Tables 1 through 3 show the results of reviewing nonmatched, unresolved and out-of-scope cases for the three areas of the test site.

As we see from these tables, clerks were able to geocode and later resolve many cases during this review. But because of the change in procedures noted above, requir-

ing confirmation of at least one additional piece of information, many previously resolved cases - mainly nonmatches - became unresolved after review. Tables 4 through 6 show that the review had the overall effect of substantially lowering the nonmatch rate in each of the three sites. The rate of unresolved cases was raised in two of the three sites. For these unresolved cases, match status was later imputed (see Diffendal, 1990).

3.1 Extended Search Results

Some movers reported Census Day addresses that could be geocoded to the block level, but could not be geocoded to or matched at a particular address. For these cases, we conducted an extended search. First we examined census questionnaire data for the entire block that we geocoded the mover(s) to. If all movers who reported that particular Census Day address could not be found and matched in that block, we examined census questionnaire data for all blocks comprising the search area as defined in section 2 of this paper. Tables 7 through 9 show the results of this extended search for the three areas of the test site.

These tables show movers who were matched at their Census Day address broken down by whether the match required searching census questionnaire data (1) at only one address (2) within the block containing the address or (3) in the surrounding blocks of the search area.

As we see, relatively few matches required searching beyond a single address, and only a minority of these required searching more than one block. However, note that clerks carried out an extended search for all movers whose Census Day addresses were geocoded to the block level, including those who wound up nonmatched or unresolved. All of these nonmatched or unresolved cases required an extended search that included all of the search area. The extent of this search, coupled with the method used for accessing census questionnaire data caused major operational problems and delays. We estimate that close to 100,000 photocopies of census questionnaires had to be accessed during mover matching (see section 3.2). For the 1990 PES, one contingency plan under consideration involves limiting the search area to the block containing the geocoded address,

thus eliminating all surrounding blocks from the search area.

3.2 Microfilm Access

In 1988, census questionnaires were not sorted by census geography. This meant that there was no practical way to access the original census questionnaires. Instead, we had to access filmed images of the questionnaires on Microfilm Access Devices (MADs) to obtain photocopies of the inside of the questionnaire which contained data needed for matching. Typically, each questionnaire in a given block had to be looked up on a separate reel of film and photocopied. A major backlog developed because of the time consuming nature of the film lookup operation and the small number of MADs (three in all) available at the start of mover matching. More MADs were later obtained, but the mover matching operation continued to be hindered by competition with another census operation in accessing reels of microfilm.

Since census questionnaires are not being sorted for 1990, the Census Bureau purchased an average of 30 MADs per Processing Office (PO) in order to address the issue of adequate access to the census questionnaire data. This is several times the amount that was in the PO in 1988, where the mover matching workload was comparable to what we expect for a typical 1990 PO. Procedures for controlling the mover matching workflow, absent in 1988, have been developed for 1990 to address both the complexity of the mover matching operation and the issue of competition in accessing reels of microfilm.

4. SUMMARY AND CONCLUSIONS

Matching movers in a PES has always been difficult. The 1988 PES was no exception. Those working on mover matching encountered difficulties both in establishing the correct census geography in which to search for a mover, and in accessing census questionnaire data for addresses in the search area for those movers remaining nonmatched. The subsequent review of nonmatched and unresolved movers seemed to verify that given better resources and procedures, more mover addresses could be geocoded, and information given

by a PES respondent regarding cross streets, neighbors and census day roster could be used in confirming that the search is being conducted in the correct area. Such improvements were incorporated into plans for matching movers in the 1990 PES.

REFERENCES

- Childers, D. and Hogan, H. (1990), "Results of the 1988 Dress Rehearsal Post-Enumeration Survey" American Statistical Association Proceedings of the Section on Survey Research Methods, to appear.
- Diffendal, G. (1990), "Missing Data in the 1988 Dress Rehearsal Post-Enumeration Survey" American Statistical Association Proceedings of the Section on Survey Research Methods, to appear.
- Fay, R.E., J. Passel, J. Robinson (1988), "The Coverage of Population and Housing in the 1980 Census," 1980 Census of Population and Housing Evaluation and Research Report, PHC80-E4. Bureau of the Census, Washington, D.C.: U.S. Department of Commerce.

Hogan, H. and Wolter K. (1988), "Measuring Accuracy in a Post-Enumeration Survey," *Survey Methodology*, 14, 99-116.

Footnotes

* 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.

¹ Tape Address Register areas are areas in which commercial mailing lists are used. These areas are typically within the most-populated urban areas of the U.S.

² Prelist areas are areas in which an independent canvassing operation is done in order to compile a mailing list. Commercial mailing lists cannot be used in these areas because of insufficient coverage or inadequate mailing addresses.

³ List/Enumerate areas are areas in which a door-to-door personal visit is conducted. These are primarily the more rural areas of the country, where commercial mailing lists cannot be used, and where it is not feasible to prelist.

Table 1. Results of Mover Matching Review
St. Louis

Initial Match	Matched	Nonmatched	Unresolved	Out-of-scope	Total
Nonmatched	4	100	61	0	165
Unresolved	5	8	18	14	45
Out-of-scope	12	0	14	0	26
Total	21	108	93	14	236

Table 2. Results of Mover Matching Review
East Central Missouri

Initial Match	Matched	Nonmatched	Unresolved	Out-of-scope	Total
Nonmatched	15	25	42	1	83
Unresolved	27	11	98	6	142
Out-of-scope	0	0	0	0	0
Total	42	36	140	7	225

Table 3. Results of Mover Matching Review
Eastern Washington State

Initial Match	Matched	Nonmatched	Unresolved	Out-of-scope	Total
Nonmatched	13	11	11	1	36
Unresolved	0	7	5	0	12
Out-of-scope	3	0	2	0	5
Total	16	18	18	1	53

Table 4. Summary of Mover Matching
St. Louis

Match Category	Before Review		After Review	
	Persons	% of Total	Persons	% of Total
Matched	206	30.4	227	33.5
Nonmatched	177	26.1	120	17.7
Unresolved	70	10.3	105	15.5
Out-of-scope	224	33.1	225	33.2
Total	677		677	

Table 5. Summary of Mover Matching
East Central Missouri

Match Category	Before Review		After Review	
	Persons	% of Total	Persons	% of Total
Matched	263	30.3	305	35.1
Nonmatched	92	10.6	45	5.2
Unresolved	184	21.2	182	21.0
Out-of-scope	329	37.9	336	38.7
Total	868		868	

Table 6. Summary of Mover Matching
Eastern Washington State

Match Category	Before Review		After Review	
	Persons	% of Total	Persons	% of Total
Matched	99	35.2	115	40.9
Nonmatched	40	14.2	22	7.8
Unresolved	21	7.5	27	9.6
Out-of-scope	121	43.1	117	41.6
Total	281		281	

Table 7. Extended Search Results
St. Louis

Required Search	Mover Matches	
	Number	Percent of Total
Single Address	208	91.6
Block	17	7.5
Surrounding Blocks	2	0.9

Table 8. Extended Search Results
East Central Missouri

Required Search	Mover Matches	
	Number	Percent of Total
Single Address	287	94.1
Block	10	3.3
Surrounding Blocks	8	2.6

Table 9. Extended Search Results
Eastern Washington

Required Search	Mover Matches	
	Number	Percent of Total
Single Address	109	94.8
Block	1	0.9
Surrounding Blocks	5	4.3