

RESULTS FROM THE INTENSIVE ADDRESS LISTING EXPERIMENT:
The Effects of Different Address Listing Procedures on Housing Coverage and Mail Response

Maria E. Urrutia-Cortez and James B. Treat, U.S. Bureau of the Census*
Maria E. Urrutia-Cortez, U.S. Bureau of the Census, SRD, Room 3211-4, Washington, DC 20233-0001

Key Words: Completion Rates, Contact, Comparisons

1. INTRODUCTION

The Census Bureau conducted an address listing experiment with two objectives. The first objective was to test whether intensive address listing procedures could locate housing units that would be missed if 1990 census address listing procedures were implemented. The goal of the intensive address listing procedures was to improve housing unit coverage in difficult to enumerate urban areas. The second objective was to test varying degrees of respondent contact during address listing. The goal was to determine the effect increased respondent contact during address listing has on mailback rates. Mailback response is measured by "completion rate" in this paper. A split panel test consisting of three panels was conducted in the Philadelphia area during the week of June 13, 1994. The first panel consisted of listing the housing units in a block with "no contact" with the respondent, i.e. canvassing and listing a block by observation. These procedures were similar to the 1990 Post Enumeration Survey address listing procedures. The second panel consisted of listing the housing units in a block with "minimum contact" with the respondent, asking one coverage question at each housing unit listed. These procedures were similar to the 1995 Integrated Coverage Measurement address listing procedures. The third panel consisted of listing the housing units in a block with "considerable contact", asking the respondent a series of probing coverage questions. In this panel, we implemented the intensive address listing procedures. For panels 2 and 3, the lister was allowed one callback in order to speak with an eligible respondent. An eligible respondent was a household member or neighbor at least 16 years of age. Since panel 1 was listed by observation, callbacks were not required. In addition, the lister was not required to speak to a respondent in panel 1.

In each of the three panels, approximately 900 units in 20 blocks were listed. To address the first objective, addresses obtained from the third panel were matched to the 1990 Decennial Census address list (for those blocks in sample and their surrounding blocks) in order to evaluate if the probing questions identified any units missed by the 1990 census. To address the second

objective, census questionnaires were mailed to the housing units listed in the three panels and completion rates were compared among the panels.

2. BACKGROUND

A workgroup was formed in March 1992 at the Census Bureau to address issues related to whole housing unit undercoverage in household surveys. The group was established to accomplish three goals:

1. Develop a list of all possible reasons for whole housing unit undercoverage.
2. Identify the major causes of whole housing unit undercoverage from the list of all possible reasons.
3. Determine which causes should receive priority for research.

The group was principally concerned with the undercoverage of minorities (Blacks and Hispanics) in difficult to enumerate urban areas since they have the highest undercoverage rates. The group compiled a list of 52 causes of whole housing unit undercoverage and ranked them in importance. Whole housing unit undercoverage occurs when all persons in a household have no chance of selection. A housing unit is a house, an apartment, a group of rooms, or a single room in which the occupant(s) lives and eats separately from other person(s) in the building and there is direct access from the outside. The access to the outside could be through a common or public hall. The group estimated the amount of survey undercoverage due to whole housing unit and within housing unit undercoverage. They concluded that within housing unit undercoverage was larger than whole housing unit undercoverage. However, approximately 35 percent of all Black undercoverage was attributed to whole housing unit misses (Shapiro, et. al., 1993). The group concluded that it would be feasible to reduce survey undercoverage resulting from missed housing units by developing and implementing intensive address listing procedures. The more thorough procedures would be tested in urban blocks with a high concentration of minority residents. Each block would be listed using the intensive address listing procedures and the addresses would be compared to the 1990 Decennial Census addresses for that block. A comparison would be made between the two sets of addresses to see if the intensive address listing procedures identified any

missed housing units. A proposal was drafted and accepted to cover the research outlined above.

In addition, this proposal was later expanded to include research used by the 1995 Census Test Integrated Coverage Measurement program. The Integrated Coverage Measurement conducts an independent listing of housing units approximately one month prior to mailing out census questionnaires to those areas. One coverage question was asked at each housing unit listed. There was a concern among Census Bureau staff that contact with the respondent prior to the census might bias their response to the census. Thus, the two objectives were merged into the one listing experiment.

Work on this project began in November of 1993. Plans were developed to list addresses in a sample of blocks using the three different listing procedures and to mail census questionnaires to those housing units listed. Mail completion rates would then be calculated and compared among the three panels. The probing coverage questions for the third panel were developed and circulated for comments within several divisions in the Census Bureau. Listing procedures and training were also developed. In May of 1994, the probing questions were tested in a difficult to enumerate urban area in Baltimore, Maryland by two experienced Senior Field Representatives. Comments from the Senior Field Representatives were also incorporated into the probing questions. In addition, the sample was identified in May of 1994.

Training for the listing experiment was conducted in the Philadelphia Regional Office on June 13, 1994. Twelve Field Representatives were trained. These Field Representatives were experienced interviewers, not experienced listers. On the morning of June 14, the trainees were divided into groups of two and given a block to list as part of their training. They were to call the supervisor with any questions after the practice listing. They started their regular assignment on the afternoon of June 14th. The listers worked independently unless there were language difficulties. In these cases, a facilitator was used to help with language problems. Most of the listings were completed by June 17, 1994. The question *How many housing units are in this building?* was completed by observation in panel 1 (no contact) and by asking an eligible respondent in panel 2 (minimum contact). Addresses were keyed at the end of July in order to create a check-in file and to generate questionnaire labels. The questionnaires were mailed on August 3rd. The 1994 Census Test short-form respondent-friendly questionnaire was mailed to the three panels. In addition, the outgoing questionnaire envelope contained

a mandatory motivational message stating "U.S. Census Form Enclosed, YOUR RESPONSE IS REQUIRED BY LAW." Questionnaire check-in took place from August 8 through September 14, 1994. The completion rates for the three panels were calculated by mid-September.

Work began on the initial match of panel 3 (considerable contact) to the 1990 Decennial Census list in August 1994. The initial match was conducted at the Jeffersonville, Indiana Processing Office. The staff performing the initial match were experienced matching specialists that had worked on the 1990 Post Enumeration Survey address matching process. Unresolved and unmatched cases were reconciled in the field during a followup operation. The followup questionnaire was designed in late August. The followup operation was started in mid-September and was completed in December 1994. The final clerical match was conducted during January and February of 1995. Data analysis was completed in March 1995.

3. METHODOLOGY

3.1 Sample Design

The Census Bureau asked the Philadelphia Regional Office to identify 60 blocks with the following characteristics based on 1990 census data:

1. Blocks located in the urban, inner-city areas of Philadelphia.

2. Blocks containing only 30-50 housing units (including vacant units).

3. Blocks containing a mixture of single and small multi-unit buildings (less than 10 units). Therefore, no blocks should contain any high rise buildings or large apartment complexes.

4. Blocks with high concentrations of a minority group. Minority is defined as Black, Hispanic or Immigrant. Black and Hispanic blocks were identified based on 1990 census block level data. Immigrant blocks were identified based on 1990 census tract level data. High concentration is defined as at least 50 percent. The following six categories (strata) were defined:

- a. Greater than 80 percent Black - 12 blocks
- b. Greater than 80 percent Hispanic - 12 blocks
- c. Greater than 80 percent Immigrant - 9 blocks
- d. 50 to 80 percent Black - 9 blocks
- e. 50 to 80 percent Hispanic - 9 blocks
- f. 50 to 80 percent Immigrant - 9 blocks

5. Approximately six of the blocks should be "run-down, drug-infested" blocks, which may be distributed in any of the strata. The Philadelphia Regional Office was not required to identify these

blocks. The purpose of this selection criteria was to guarantee that the worst of the hard to enumerate blocks would be included in the experiment.

The objective of the experiment was to look for large difference in the completion rates, therefore, the total sample size was small, 60 blocks. The Census Bureau requested that the Philadelphia Regional Office identify the 60 sample blocks and at least 6 additional blocks to be used during the training. The Philadelphia Regional Office identified 72 blocks meeting the above criteria.

After obtaining the list of blocks by stratum from the Philadelphia Regional Office, the Census Bureau allocated the sample blocks within a stratum to a panel. Within each stratum, blocks were clustered based on their geographic proximity (census tract and block number). Each cluster consisted of three blocks. Within each cluster, blocks were randomly assigned to a panel.

A total of 2,789 housing units were listed in the experiment. In the "no contact", the "minimum contact" and the "considerable contact" panels, 905, 1,026 and 858 housing units were listed, respectively.

3.2 Definition of a Postmaster Return

For this experiment, a case was considered a postmaster return if the questionnaire was returned by the United States Postal Service as an undeliverable.

3.3 Add Rate

The add rate was calculated using the following formula:

$$\text{Add Rate} = \frac{\text{Correct PALE Addresses} - \text{Correct 1990 Census Addresses}}{\text{Correct 1990 Census Addresses}} * 100$$

PALE represents the Philadelphia Address Listing Experiment

3.4 Mail Completion Rate

The mail completion rates were calculated for the three panels using the following formula:

$$\text{Mail Completion Rate} = \frac{\text{Total Mail Returns}}{\text{Total Mailed} - \text{Postmaster Returns}}$$

The mail completion rates are based on mail returns and postmaster returns received between August 3 and

September 14, 1994.

3.5 Dual-System Estimator

The usual dual-system estimator is used to estimate the number of housing units missed by both operations; the 1990 census and the Philadelphia Address Listing Experiment. The assumption is that there is independence between the two listing operations.

Given:

Philadelphia Address Listing Experiment	1990 Decennial Census		
	Listed	Not Listed	Total
Listed	$N_{1,1}$	$N_{1,2}$	$N_{1..}$
Not Listed	$N_{2,1}$	$N_{2,2}$	$N_{2..}$
Total	$N_{.,1}$	$N_{.,2}$	$N_{...}$

Then:

$$\text{Dual-System Estimator} = \frac{N_{2,1} * N_{1,2}}{N_{1,1}}$$

3.6 Statistical Inference

The mail completion rates were calculated for each panel. The stratum level completion rates are not reported due to the small sample size. The standard errors for the panel estimates were computed using the Stratified Jackknife variance procedure (Wolter, 1985) with the unit being the block mail completion rate. The estimates were produced by the VPLX statistical software.

4. RESULTS

4.1 Mail Response

The mail completion rates for the three panels are presented in Table 1. The experiment focused on high crime and high minority areas, which in the past have had low mail response. Therefore, we anticipated low completion rates for this experiment. For the panel with no contact during the address listing, the mail completion rate was 25.4 percent. The mail completion rate for the panel with the minimum contact during address listing was 20.3 percent. Finally, the mail

completion rate for the third panel (considerable contact) was 24.7 percent.

Table 2 summarizes each of the paired comparisons between the three different address listing procedures. The first comparison is panel 1 (no contact) with panel 2 (minimum contact). The second comparison is panel 1 (no contact) with panel 3 (considerable contact). Finally, the last comparison is panel 2 (minimum contact) with panel 3 (considerable contact). For all comparisons the differences are not statistically significant at the 90 percent confidence interval. Therefore, there is no evidence of a difference among the different address listing procedures on completion rates. On the other hand we have not proven there is not a difference in the completion rates. Based on the observed data, the difference between panels 1 and 2 could be as large as 11.4 percent, with no evidence of a difference. Similarly, the difference between panels 1 and 3 and panels 2 and 3 could be as large as 5.9 percent and 10.3 percent, respectively, with no evidence of a difference. Note that the observed estimates do not follow any expected pattern. Therefore, the differences between the panels are most likely due to random variations and not due to the different address listing procedures. The strata differences were not reported due to the small sample size.

4.2 Housing Coverage

Table 3 contains the comparison of the address lists from the 1990 Decennial Census and the Philadelphia Address Listing Experiment (PALE) for the 20 blocks in panel 3. During the 1990 Decennial Census, the Census Bureau identified 871 addresses. In June of 1994, the Philadelphia Address Listing Experiment identified 858 addresses for the same geographic area. From the address matching operation, 749 addresses were identified on both the 1990 census and the Philadelphia Address Listing Experiment address lists. This represented approximately 86 percent and 87 percent of the addresses from the 1990 census and the Philadelphia Address Listing Experiment, respectively. The 1990 census address list had an additional 122 addresses which were not listed in the Philadelphia Address Listing Experiment. Similarly, the Philadelphia Address Listing Experiment address list contained 109 addresses which were not in the 1990 Decennial Census. Using the usual dual-system estimator and assuming independence between the 1990 census and the Philadelphia Address Listing Experiment, the expected number of housing units missed by both operations is 18. During the field followup operation for the address matching operation,

four addresses were identified which were not on either the 1990 census or the Philadelphia Address Listing Experiment address lists.

The numbers in Table 3 represent only half of the picture on coverage. Table 4 contains a breakdown of the address status by the operation (1990 census and/or Philadelphia Address Listing Experiment) which listed the address. The address status was determined during the address matching and field followup operations. The addresses were classified as correct, erroneous or undetermined. The Jeffersonville Processing Office sent 203 (20.6 percent) of the 984 addresses to field followup.

For the 749 addresses listed during both the 1990 census and the Philadelphia Address Listing Experiment, 733 addresses (97.9 percent) and 5 addresses (0.7 percent) were correct and erroneous, respectively. In addition, the address status for the remaining 11 addresses (1.5 percent) was undetermined.

For the 122 addresses listed only during the 1990 census, only 12 addresses (9.8 percent) were correct. The field followup enumerators were able to locate these 12 addresses and verify that they were missed during the Philadelphia Address Listing Experiment. The census address list contained 102 addresses (83.6 percent) which were erroneous. All 102 erroneous addresses were located during the field followup operation. For 55 of the 102 addresses, the field followup enumerators were able to locate the basic street address and unable to locate the housing unit within the basic street address. In addition, the field followup enumerators indicated that there had been no changes to the structure since April 1, 1990. Therefore, these 55 addresses were "true" erroneous enumerations in the 1990 census. Forty-four of the 102 erroneous addresses were listed in the wrong block during the 1990 census; geocoding errors in the 1990 census. The large number of geocoding errors was not surprising since, during the 1990 census the Census Bureau had difficulty geocoding addresses in the Philadelphia area. The matching clerks and/or field followup enumerators located 39 of these 44 addresses in the surrounding blocks. For three of the 102 erroneous addresses, the field followup enumerators located the structure but were unable to determine the status of the address on April 1, 1990. Two addresses were in a structure that was boarded up. The field followup enumerator was unable to locate anyone. For the remaining one address which was classified as erroneous, the field followup operation obtained no information on the status of the address in 1990. Finally, the status of the remaining eight addresses (6.6 percent) listed only during the 1990 census was undetermined. The field followup enumerators were

unable to locate these eight addresses. They could have existed during the 1990 census and been demolished prior to the Philadelphia Address Listing Experiment. Therefore, these eight addresses do not necessarily represent erroneous enumerations in the 1990 census.

For the 109 addresses listed only during the Philadelphia Address Listing Experiment, 46 addresses were correct based on the field followup operation. This represented 42.2 percent of the addresses only listed in the Philadelphia Address Listing Experiment. We matched the addresses to the surrounding blocks and determined the addresses were not just geocoding errors in the 1990 census. The 46 addresses were not located in the block or the surrounding blocks in the census list. The field followup operation verified these addresses existed on April 1, 1990, therefore they are true misses in the 1990 census. The remaining 51 addresses (46.8 percent) and 12 addresses (11.0 percent) were either erroneous or the address status was undetermined, respectively.

Finally, the four addresses which were identified during the field followup operation as missed in both the 1990 census and the Philadelphia Address Listing Experiment were actual addresses. Therefore, these addresses were classified as correct.

Based on the results from Table 4, the 1990 census address list contained 745 correct addresses (733 and 12). Similarly, the Philadelphia Address Listing Experiment address list identified 779 correct addresses (733 and 46). Therefore, the add rate is approximately 4.6 percent. If we consider the undetermined addresses as correct, then the 1990 census and the Philadelphia Address Listing Experiment address lists identified 764 (733, 12, 11 and 8) and 802 (733, 46, 11 and 12) correct addresses, respectively. Therefore, the add rate would increase slightly to approximately 5.0 percent.

5. CONCLUSION

The results from the Philadelphia Address Listing Experiment contain two main findings. First, there was no evidence of a difference in the completion rates among the three panels. Second, the analysis of the addresses listed using the considerable contact address listing procedures (panel 3) resulted in added addresses, approximately 4.6 percent to 5.0 percent. Based on our findings, we recommend that additional research be conducted using intensive address listing procedures. Specifically, the same address matching and field followup operations used in Panel 3 should be conducted for Panels 1 and 2. A comparison of the add

rates among the three panels would complete the analysis. The panel with the largest add rate would indicate the address listing procedure which would produce the best address list. If the add rates were the same among the three panels, then the panel which is the easiest to implement (panel 1 - no contact) should be used. A potential problem of completing the analysis now, is that the field followup operation would take place more than a year after the original operation and this might cause discrepancies in the data due to the time difference.

In conclusion, we recommend that research continue to find additional ways of reducing whole housing unit undercoverage, specifically in mobile homes, special places and new construction (Shapiro, et. al., 1993).

6. REFERENCE

Shapiro, Gary, Chip Alexander, Larry Cahoon, Gregg Diffendal, Sheila Grimm, Howard Hogan, Cliff Loudermilk, and Barbara Williams, 1993, "Final Report of the Whole Unit Undercoverage Group", Unpublished Report, U.S. Bureau of the Census, August 25, Draft.

Wolter, Kirk, 1985, Introduction to Variance Estimation. New York, Springer-Verlag.

7. ACKNOWLEDGEMENTS

The authors would like to thank the following persons for their contributions to the implementation of the Philadelphia Address Listing Experiment. Jan Jaworski and David Pence of the Field Division for the preparation of the lister's instructions and the liaison work with the Philadelphia Regional Office, respectively. Charisse Jones of the Statistical Research Division for supervising the keying of the addresses and creating the computer files for the experiment. Judith Dawson, Dennis Stoudt and staff of the Decennial Management Division for questionnaire processing. Martha Sutt and her staff (Lynn Evans, Estell Power and Vicki Smith) of the Data Preparation Division for conducting the address matching operation. Howard Hogan of the Services Division, Jeffrey Moore and Raj Singh of Statistical Research Division for their review of the paper.

* 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.

Table 1: Philadelphia Address Listing Experiment Mail Completion Rates by Panel

Panel	Mail Completion Rate Estimates (%) and Standard Errors (%)	
	Estimate	Standard Error
1. No Contact	25.4	2.2
2. Minimum Contact	20.3	2.9
3. Considerable Contact	24.7	2.0

Table 2: Comparisons Between the Different Address Listing Procedures

Experimental Comparisons	Completion Rate Differences (%), Standard Error (%) and 90% Confidence Intervals (C.I.)		
	Difference	Standard Error	90% C.I.
1 - 2	5.1	3.8	-1.2 to 11.4
1 - 3	0.7	3.2	-4.5 to 5.9
2 - 3	-4.4	3.6	-10.3 to 1.5

Table 3: Comparison of the Address Lists from the 1990 Decennial Census and the Philadelphia Address Listing Experiment

Philadelphia Address Listing Experiment	1990 Decennial Census		
	Listed	Not Listed	Total
Listed	749	109	858
Not Listed	122	4	126
Total	871	113	984

Table 4: A Breakdown of Address Status by the Operation where the Addresses were Listed

Address Status	The Operation where the Addresses were Listed			
	Both	Census Only	PALE Only	Neither
Total:	749	122	109	4
Correct	733	12	46	4
Erroneous	5	102	51	0
Undetermined	11	8	12	0