

Analysis of Proxy Data in the Accuracy and Coverage Evaluation

FINAL REPORT

This evaluation study reports the results of research and analysis undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning. Findings from the Census 2000 TXE Program reports are integrated into topic reports that provide context and background for broader interpretation of results.

Glenn Wolfgang,
Rosemary Byrne, and
Shermaine Spratt

Decennial Statistical
Studies Division

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

CONTENTS

EXECUTIVE SUMMARY	v
1. BACKGROUND	1
1.1 What was the role of Proxy data collection in the Census 2000 A.C.E.?	1
1.2 How were Proxy data collected in the A.C.E. Person Interview?	1
1.3 How were Proxy data collected in the A.C.E. Person Followup Interview?	2
1.4 What other reports provided background to the quality of proxy data?	3
2. METHODS	3
2.1 How were results analyzed?	3
3. LIMITS	4
4. RESULTS	5
4.1 What were the percentages of proxy enumerations for various data conditions? ...	5
4.2 What variables were related to the percent of proxy reporting?	9
4.3 How did the percent of proxy enumeration relate to consistency between matched Census 2000 and A.C.E. data?	22
4.4 Did proxy response on the original A.C.E. interview relate to proxy response in the followup interview	24
4.5 Did the percent of Item Non-response in the original A.C.E. interview relate to the source of proxy response?	26
4.6 Did the percent of Item Non-response in the original A.C.E. interview relate to the mover status?	30
5. CONCLUSIONS AND RECOMMENDATIONS	36
REFERENCES	38
APPENDIX: Percent of Item Non-response for Proxy versus Non-Proxy	39

LIST OF TABLES

Table 4.1.1. Percentages of Persons Reported by Proxy for Alternative Interview Operations, Reference Days, and Weights Applied	7
Table 4.1.2. Percentages of Housing Units Reported by Proxy for Alternative Interview Operations, Reference Days, and Weights Applied	8
Table 4.2.1. Percent of Original Interview Census Day Person Proxy, by Age and Sex	10
Table 4.2.2. Percent of Followup Interview Census Day Person Proxy, by Age and Sex . . .	11
Table 4.2.3. Percent of Original Interview Census Day Person Proxy, by Tenure	11
Table 4.2.4. Percent of Followup Interview Census Day Person Proxy, by Tenure	11
Table 4.2.5. Percent of Original Interview Census Day Person Proxy, by Race and Hispanic Origin	12
Table 4.2.6. Percent of Followup Interview Census Day Person Proxy, by Race and Hispanic Origin	12
Table 4.2.7. Percent of Original Interview Census Day Person Proxy, by Size of Metropolitan Statistical Area and Type of Enumeration Area	13
Table 4.2.8. Percent of Followup Interview Census Day Person Proxy, by Size of Metropolitan Statistical Area and Type of Enumeration Area	13
Table 4.2.9. Percent of Original Interview Census Day Person Proxy, by Return Rate Indicator	14
Table 4.2.10. Percent of Followup Interview Census Day Person Proxy, by Return Rate Indicator	14
Table 4.2.11. Percent of Original Interview Census Day Person Proxy, by Region of the United States	15
Table 4.2.12. Percent of Followup Interview Census Day Person Proxy, by Region of the United States	15
Table 4.2.13. Percent of Original Interview Census Day Person Proxy, by Imputation of Characteristics	16
Table 4.2.14. Percent of Followup Interview Census Day Person Proxy, by Imputation of Characteristics	16
Table 4.2.15. Percent of Original Interview Census Day Person Proxy, by Type of Structure at Basic Street Address	16
Table 4.2.16. Percent of Followup Interview Census Day Person Proxy, by Type of Structure at Basic Street Address	17
Table 4.2.17. Percent of Original Interview Census Day Person Proxy, by TES Status	17
Table 4.2.18. Percent of Followup Interview Census Day Person Proxy, by TES Status	17
Table 4.2.19. Percent of Original Interview Census Day Person Proxy, by Targeted Extended Search and Relist Cluster Status	18
Table 4.2.20. Percent of Followup Interview Census Day Person Proxy, by Targeted Extended Search and Relist Cluster Status	18
Table 4.2.21. Percent of Original Interview Census Day Person Proxy, by Match Status	19
Table 4.2.22. Percent of Followup Interview Census Day Person Proxy, by Match Status . . .	19
Table 4.2.23. Percent of Original Interview Census Day Person Proxy, by Mover Status	19

Table 4.2.24. Percent of Followup Interview Census Day Person Proxy, by Mover Status . .	19
Table 4.2.25. Percent of Original Interview Census Day Person Proxy, by Household Size	20
Table 4.2.26. Percent of Followup Interview Census Day Person Proxy, by Household Size	20
Table 4.2.27. Percent of Original Interview Census Day Person Proxy, by Kinship to Reference Person	21
Table 4.2.28. Percent of Followup Interview Census Day Person Proxy, by Kinship to Reference Person	21
Table 4.2.29. Percent of Original Interview Census Day Person Proxy, by Sample Exclusion Status	22
Table 4.2.30. Percent of Followup Interview Census Day Person Proxy, by Sample Exclusion Status	22
Table 4.3.1. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Age and Sex Categories	23
Table 4.3.2. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Tenure Categories	23
Table 4.3.3. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Race and Hispanic Origin Domain Categories	23
Table 4.3.4. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Type of Structure at Basic Street Address Categories	23
Table 4.3.5. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Household Size Categories	24
Table 4.4.1. Frequencies and Percents of Persons with Proxy Report: Original Interview by Followup Interview	24
Table 4.4.2. Percent of Persons By Proxy Status Combination within Tenure Groups	25
Table 4.4.3. Percent of Persons By Proxy Status Combination within Race and Hispanic Origin Domain Groups	25
Table 4.4.4. Percent of Persons By Proxy Status Combination within Age and Sex Groups	26
Table 4.5.1. Percent of Item Non-response to the Tenure Item, by Type of Proxy Respondent	27
Table 4.5.2. Percent of Item Non-response to the Age Item, by Type of Proxy Respondent	27
Table 4.5.3. Percent of Item Non-response to the Gender Item, by Type of Proxy Respondent	28
Table 4.5.4. Percent of Item Non-response to the Race Item, by Type of Proxy Respondent	28
Table 4.5.5. Percent of Item Non-response to the Hispanic Origin Item, by Type of Proxy Respondent	29
Table 4.5.6. Percent of Item Non-response to the Kinship to Reference Person Item, by Type of Proxy Respondent	29
Table 4.5.7. Percent of Item Non-response to the Other Residence Flag Item, by Type of Proxy Respondent	30

Table 4.5.8.	Percent of Item Non-response to the Special Place Flag, by Type of Proxy Respondent	30
Table 4.6.1.	Percent of Item Non-response to the Tenure Item, by Mover Status	31
Table 4.6.2.	Percent of Item Non-response to the Age Item, by Mover Status	32
Table 4.6.3.	Percent of Item Non-response to the Gender Item, by Mover Status	32
Table 4.6.4.	Percent of Item Non-response to the Race Item, by Mover Status	33
Table 4.6.5.	Percent of Item Non-response to the Hispanic Origin Item, by Mover Status	33
Table 4.6.6.	Percent of Item Non-response to the Kinship Item, by Mover Status	34
Table 4.6.7.	Percent of Item Non-response to the Other Residence Item, by Mover Status	34
Table 4.6.8.	Percent of Item Non-response to the Special Place Flag Item, by Mover Status	35
Table A.1.	Percent of Item Non-response to the Tenure Item, by Proxy Status	39
Table A.2.	Percent of Item Non-response to the Age Item, by Proxy Status	39
Table A.3.	Percent of Item Non-response to the Gender Item, by Proxy Status	39
Table A.4.	Percent of Item Non-response to the Race Item, by Proxy Status	39
Table A.5.	Percent of Item Non-response to the Hispanic Origin Item, by Proxy Status ..	40
Table A.6.	Percent of Item Non-response to the Kinship Item, by Proxy Status	40
Table A.7.	Percent of Item Non-response to the Other Residence Flag Item, by Proxy Status	40
Table A.8.	Percent of Item Non-response to the Special Place Flag Item, by Proxy Status ..	40

EXECUTIVE SUMMARY

The goal of the Analysis of Proxy Data in the Accuracy and Coverage Evaluation P sample was to identify characteristics of persons or households that may be related to whether the interview respondent was a household member or a proxy, that is, some other respondent knowledgeable about the household. Understanding the conditions surrounding proxy response is essential to designing and evaluating data collection for coverage measurement.

We computed alternatively defined percents of proxy response. Definitions differed regarding reference to Census Day versus Interview Day, original versus followup interview, person versus housing unit data, and alternative estimation weights. We compared percents of proxy response for groups formed from variables relevant to Accuracy and Coverage Evaluation operations, like poststratification, including, for example, age, sex, tenure, imputation status, and mover status. We also analyzed percents of proxy response by consistency, from census to Accuracy and Coverage Evaluation, in answers to some of those same variables. We cross-tabulated proxy versus non-proxy response in the original interview with proxy versus non-proxy in the followup for all followup persons as well as different groups based on age, sex, tenure, race, and Hispanic origin. We analyzed percents of item non-response to selected variables by type of respondent (non-proxy, landlord, neighbor, postal worker, observation or other) and by mover status.

Did any results point to particular groups of persons having a higher percent of proxy response?

Yes. The overall percentage of proxy respondents for weighted Census Day persons was 5.5 percent for the original interview and 32.4 percent for the followup interview. The higher percentage of proxy for followup interviews is explained by the selection of followup cases largely upon proxy status in the original interview. Groups of persons that had percentages of proxy respondents significantly higher than other groups included the following. (The first percent shown refers to the Accuracy and Coverage Evaluation original person interview and the second percent shown refers to the person followup interview, unless otherwise noted.)

Age and Sex Groups:

- Young adult females aged 18-29 (8.0 and 37.5 percents),
- Young adult males aged 18-29 (7.2 and 37.0 percents),
- Adult males aged 30-49 (6.2 and 36.9 percents),
- Adult females aged 30-49 (5.7 and 34.3 percents),

Tenure and Residence Groups:

- Non-owners (10.5 and 45.1 percents),
- Persons in multi-unit dwellings (12.0 and 46.1 percents),
- Outmovers, in comparison to nonmovers (74.8 and 70.3 percents),
- Persons living alone (13.5 and 57.9 percents),

Geographic Groups:

- Persons in designated low census-return-rate areas (6.5 percent – original interview only),
- Persons in South (6.1 and 33.3 percents) and West (5.9 and 34.3 percents) regions,

Characteristics and Processing Status Groups:

- Persons with imputed characteristics (21.6 and 41.0 percents),
- Targeted Extended Search persons (8.4 percent – original Accuracy and Coverage Evaluation interview only), and
- Persons with imputed match status, i.e., unresolved search for a matching census enumeration (55.1 and 50.1 percents.)

Also, census-matched persons for whom answers to the original Accuracy and Coverage Evaluation interview questions were not the same as the census had higher percentages of proxy respondents. Those percentages ranged from 6.2 to 13.5 percent among items analyzed.

Conversely, some household characteristic groups had statistically significant lower percentages of proxy response relative to other groups, including the following (Percents for the original and followup interviews are listed in parentheses as above.):

- Parent or child of the person listed as reference person (3.8 and 24.4 percents) and
- Persons in households of seven or more residents (2.5 and 16.4 percents.)

Did any results point to particular types of respondents yielding a higher or lower quality of response as measured by item non-response on selected data items?

Yes, responses provided by household members rather than proxies had the lowest percents of item non-response. In the other extreme, postal workers responding as proxies yielded the highest rates of item non-response to many important questions, but not to race and Hispanic origin items, for which landlords provided the highest percents of item non-response.

Did any results point to particular mover status groups having a higher or lower quality of response as measured by item non-response on selected data items?

Not in general, the percents of item non-response from proxy respondents for whole-household in-movers, whole-household out-movers, and non-movers were not consistent across items analyzed. Out-movers from non-mover households had the highest percents of item non-response to race and Hispanic origin questions. Whole-household movers, both in-movers and out-movers, had higher percents of item non-response than non-movers for race. The opposite was true for tenure. No general pattern of item non-response among mover groups was evident.

What were the implications of these results?

Proxy responses are associated with important effects in coverage measurement results. Understanding those effects is important to interpreting past coverage measurements and planning future ones. In particular, we observed in many analyses that proxy response was related to unresolved or less complete answers to important coverage measurement questions. Although we cannot say that proxy response is the cause of less informed data, we recommend continuing efforts to minimize the early acceptance of proxy response as a guideline in data collection. Continuing to evaluate proxy data in future data collections would be helpful to replicate or extend these findings. We recommend pursuing alternatives to proxy data collection, such as use of administrative records, to potentially yield more accurate data.

1. BACKGROUND

This report summarizes proxy data collection procedures and analyses concerning proxy effects on the P sample of the Census 2000 Accuracy and Coverage Evaluation (A.C.E.)

1.1 What was the role of Proxy data collection in the Census 2000 A.C.E.?

The Census 2000 A.C.E. measured Census 2000 coverage of housing units and the population. The A.C.E. P sample in particular gathered information to determine how many persons were missed in Census enumeration. To assess coverage, we designed the A.C.E. to obtain information about residents of a housing unit on Census Day and on the day of the A.C.E. interview. We asked the A.C.E. field staff to get information about each household from a household member, particularly enough information to list household members and later determine whether they matched anyone enumerated in Census 2000 and who was resident on Census Day. If they could not, we allowed a proxy interview, which means the A.C.E. interviewers attempted to find and get a response from a knowledgeable respondent who was not a member of the household. A knowledgeable respondent was someone who knew about the names, ages, and current or Census Day residence status of people living in the household. Neighbors, landlords, and postal workers were typical proxy respondents.

Proxy respondents supplied data during the original A.C.E. person interviews and the A.C.E. person followup (FU) interviews. During person interviewing, the A.C.E. collected names, demographic characteristics, and residency information for the people who lived at a sample housing unit on Census Day, as well as on the day of the interview. The results of person interviewing were matched to the census. The followup interview collected additional information when needed to complete determination of match and residence status.

Census Day resident data were essential in dual system estimation computations, so when the whole household of Census Day residents moved out before they could respond to the A.C.E. interview, anyone who responded, current resident or not, was considered a proxy. Census Day residents in a household where everyone moved out by the day of the A.C.E. interview were called whole-household outmovers. A.C.E. residents in a household where everyone moved in after Census Day were called whole-household inmovers.

1.2 How were Proxy data collected in the A.C.E. Person Interview?

The A.C.E. person interviews were conducted between April 2000 and September 2000 using a Computer Assisted Personal Interviewing (CAPI) instrument on a laptop computer. The interviews gathered important coverage measurement information about current (known as Interview Day) household residents, as well as Census Day (April 1, 2000) household residents. When the Census Day and Interview Day households were not the same, the data may have been provided by different respondents, one or both possibly proxy.

The A.C.E. had specific rules about when the interviewers could attempt a proxy interview. Interviewers tried to contact members of the household during the first three weeks after assignment. A proxy interview could be taken without further attempts to locate a household

member if no household member would be available for the duration of the interviewer's assignment or if a knowledgeable non-household member was the only person capable of answering the questions or after the first three weeks were past. If the interviewer could not get an interview, proxy or otherwise, in the first six weeks after assignment, a more experienced interviewer attempted to get an interview with either a household member or a proxy respondent in an operation called non-response conversion during the final two weeks of interviewing in that cluster.

The CAPI instrument for A.C.E. person interviewing had three separate paths. The main path, called the non-proxy path, was for respondents who were members of the current household. The second path, called the proxy path, was for respondents who served as proxy for the current household. These first two paths captured data for persons who lived at the sample address on both Census Day and Interview Day (nonmovers) or those who moved in after Census Day (inmovers.) A third path, called the outmover path, captured data for whole households that moved out after Census Day (outmovers.) The wording and order of the questions differed slightly in each path according to the situation.

There was no intentional tracing of outmovers. In the Census 2000 Dress Rehearsal, outmover tracing was tested (Killion, 1999) and was not found to significantly improve dual system estimates relative to proxy reporting. All data collected in the outmover path of the CAPI instrument were considered proxy, although sometimes a neighbor or another available respondent was indeed a Census Day member of the household. In proxy interviews, the CAPI instrument asked for respondent type. The answer categories were: landlord, neighbor, postal worker, interviewer observation, and other-(specify.) All persons enumerated in a proxy A.C.E. interview and not matched to Census 2000 enumerations were sent to the person followup interview.

1.3 How were Proxy data collected in the A.C.E. Person Followup Interview?

The A.C.E. person followup interview aimed to get any additional information needed to accurately code match and residence status for P-sample people, persons enumerated by A.C.E. in sample areas to determine who was missed in Census 2000. For nonmatches, the decision to send a case to followup was based in part on whether or not the A.C.E. person interview was a proxy or non-proxy. Childers (2001) described which P-sample cases were sent to followup. In summary, persons falling in one or more of the following categories were followed up:

- Proxy response nonmatches – P-sample nonmatches in households where the data were collected during an interview with a proxy respondent.
- Partial household nonmatches – P-sample nonmatches in households where other persons matched.
- Conflicting households – P-sample nonmatches in households where housing units matched, but none of the persons did.
- Unresolved persons – Possible matches or unresolved residence status persons.

During A.C.E. person followup interviewing, interviewers generally had to make six attempts at different times and on different days to find a household member before they could try to get a proxy interview. A proxy interview could be taken without further attempts to locate a household member if no household member would be available for the duration of the interviewer's assignment or if a non-household member was the only person capable of answering the questions.

Unlike A.C.E. person interviewing, A.C.E. person followup used a paper form. It was produced by a process (Docuprint) that printed only the appropriate sections of the questionnaire. The A.C.E. person followup form had check boxes for type of respondent: household member, neighbor, relative, friend, apartment manager, landlord, care giver, and other.

1.4 What other reports provided background to the quality of proxy data?

Other reports provided greater detail on the A.C.E. and prior census coverage evaluations. Hogan (1993) reported on both analyses and procedures for the 1990 census. Childers (2001) described the A.C.E. design. Adams, Barrett, and Byrne (2001) summarized procedures for A.C.E. operations.

Past research debates the quality of proxy response. Griffin and Moriarity (1992) concluded there were fewer census coverage errors on returns from household members that return their questionnaires by mail relative to proxy returns. Hill (1987) found "substantial differential reporting bias and error variance between self and proxy reports of earnings" in the Survey of Income and Program Participation. O'Muircheartaigh (1986) and Moore (1988) noted that such studies generally do not control assignment of respondents to the proxy condition. While we should be cautious of viewing proxy response as an independently acting cause of diminished data quality, the relation of proxy response to quality indicators in live data continues to be of interest and value.

2. METHODS

2.1 How were results analyzed?

Many analyses in this study compared percentages of proxies. Those percentages were computed within subgroups of the P sample based on characteristics, like age and sex, or descriptive variables, like region. The numerator of this percent was the number of proxies in the subgroup; the denominator for the percent was the number of P-sample persons in the subgroup. In the analyses of this report's last two sections, methods similar to those described here compared percentages of non-response to interview items for subgroups of proxy respondents.

The percentages of proxies for different subgroups were compared using stratified jackknife variance estimation and pair-wise t-values generated by VPLX (Fay, 1990.) Variances were confirmed by an alternate program.

Statistical significance for these t-values was determined using the Bonferroni multiple comparison of means technique. It controlled the probability of Type I error for a family of tests. In the context of this analysis, a family of tests was defined as all tests conducted among sample subgroups formed from the variable under analysis. For example, when comparing four subgroups, six pairs of statistics were tested. To control the chance of Type I error at $\alpha = 0.10$ for all six tests combined, we used an adjusted criterion t-value associated with a joint probability equal to 0.10 that one or more of six two-tailed tests was in error. In addition, tests with subgroups based on less than 100 person records were avoided, either through collapsing with other levels or simply by dropping the level from that family of tests.

3. LIMITS

There are certain limitations in the results presented in this paper. Several are computational shortcuts that permitted the efficiency and versatility needed to conduct a wide range of analyses.

- These analyses did not analyze proxy response in Census data, including E-sample persons or households.
- Both the A.C.E. person interviews and person followup interviews contained several variables indicating a proxy interview. In some cases these variables were inconsistent in indicating proxy status. In this analysis we used a combination of the most consistent variables to identify proxy interviews.
- These analyses attempted to assess, in a number of different ways, differences associated with proxy response. We cannot control whether or not the interview was conducted with a proxy; the households are not randomly assigned to a proxy or non-proxy response. This means that we cannot determine whether proxy response may be the reason for differences in the data. Observed differences in data quality may be due to something innate about the household, which also influences who the respondent was.
- Percent and standard error computations in the proxy evaluation analyses were simplified and did not take into account all aspects of the sampling and estimation, which were technically problematic under the subgrouping conditions. Official estimates are generated by summing (to the national or other levels) computations within designed poststrata. Within those poststrata, under certain data conditions, in-mover data were used in place of out-mover data in estimation. For this evaluation, estimates were computed without regard to the original poststrata, as if all records were in a single cell within the analysis group, and only non-mover and out-mover data were used. Also, missing data imputation variance was ignored.

4. RESULTS

Proxy rates were first computed across all records for a variety of conditions. The different results showed how different percent-proxy statistics would be with differently defined reference groups. They also provided a baseline or context for subgroup statistics of the same type. Other analyses explored the following research questions:

- Did the percent of proxy responses relate to variables such as region, home ownership or race?
- Did the percent of proxy responses relate to consistency in response between A.C.E. and matched census reports computed within characteristics such as region or race?
- Did proxy response on the original A.C.E. interview relate to proxy response in the A.C.E. followup interview?
- Did the percent of “Don’t Know” responses among proxies relate to the source of proxy response (e.g., landlord, postal worker)?
- Did the percent of “Don’t Know” responses among proxies relate to mover status?

4.1 What were the percentages of proxy enumerations for various data conditions?

All percent of proxy statistics considered here were computed as the percent of proxy interviews among all interviews, ignoring vacant unit and noninterview records. Data may differ in a number of other ways relevant to the definition of a percent of proxy statistic:

- In which A.C.E. operation were the data collected?
 - original person interview
 - followup of person data
- What was the reference day (the day which the percent describes)?
 - Census Day
 - A.C.E. Interview Day
- What was the unit of analysis?
 - person-level record
 - housing unit record
- How were the data weighted?
 - weighted fully for estimation
 - weighted for sample cluster, after sample reduction and large-block subsampling
 - unweighted

The percent of proxy referring to Census Day original interview persons weighted for estimation was the statistic used most often in the analyses following. It may serve as a benchmark to which other estimates may be compared. The A.C.E. Interview Day data were distinguished from Census Day data by in-mover records which are used in estimation in only certain conditions (Haines 2001). Followup data were worth exploring because they are very important to final codes for the subset of cases that needed followup. Housing unit statistics are of general interest, but many variables could be investigated only at the person level.

Estimation weights for Census Day data differed from those for Interview Day data. In dual system estimate computations, Census Day data were adjusted for noninterviews among nonmovers and outmovers, and sampling for the Targeted Extended Search operation (designed to deal with geocoding error), in addition to all cluster sampling operations prior to interviewing. Interview Day data were weighted to adjust for noninterviews among nonmovers and inmovers and for all cluster sampling. (Haines, 2001.) Official housing unit dual system estimates were not computed from person interview data, so there were no comparable estimation weights for housing unit proxy percents, but cluster weights (specifically PWGHT, the cluster weight for anything that made it into the P sample, including sample reduction and large-block subsampling operations) were used for housing unit statistics.

Tables 4.1.1 and 4.1.2 present the values of a variety of alternate percent of proxy statistics computed from person or housing unit records for the different types of interview operations, reference day, and sample weighting. Statistical tests discussed below were done using VPLX and accounting for the correlations between the different proxy statistics.

The statistic reported on the first line of Table 4.1.1 is the percent of proxy for estimation-weighted Census Day persons in the original interview (one hundred times the estimation-weighted number of original interview Census Day persons captured in a proxy response and divided by the estimation-weighted total number of original interview Census Day persons in all responses.) The next two lines of the table were computed from sample-cluster-weighted data and unweighted data, respectively. The next part of the table has followup Census Day proxy percents, starting with the estimation-weighted one (one hundred times the estimation-weighted number of followup interview Census Day persons captured in a known proxy response and divided by the estimation-weighted total number of followup Census Day persons with known proxy status.) In the next part of the table, “Combined” implies “in either the original or followup interview.” A complication in computing these percents was the 15 percent of followup cases that had unknown proxy status. Those cases were ignored in computations and the cases with known proxy status were multiplied, in the first percent of this set, by the inverse estimation-weighted proportion of usable data (one hundred times the sum of estimation-weighted original interview Census Day persons captured in a proxy response plus the number of followup interview Census Day persons captured in a known proxy response but inflated for unknown proxy response and divided by the estimation-weighted total number of original interview Census Day persons.) The remaining parts of the table are defined by substituting “Interview Day” for “Census Day.”

Table 4.1.1. Percentages of Persons Reported by Proxy for Alternative Interview Operations, Reference Days, and Weights Applied

Operation	Reference Day	Weights Applied	Percent of Proxy	s.e.
Original	Census Day	Estimation	5.5	0.1
Original	Census Day	Sample Cluster	5.5	0.1
Original	Census Day	none	5.7	0.1
Followup	Census Day	Estimation	32.4	0.6
Followup	Census Day	Sample Cluster	32.0	0.5
Followup	Census Day	none	31.4	0.5
Combined	Census Day	Estimation	6.3	0.1
Combined	Census Day	Sample Cluster	6.5	0.1
Combined	Census Day	none	6.9	0.1
Original	Interview Day	Estimation *	4.5	0.1
Original	Interview Day	Sample Cluster	4.5	0.1
Original	Interview Day	none	4.7	0.1
Followup	Interview Day	Estimation *	29.4	0.5
Followup	Interview Day	Sample Cluster	29.3	0.5
Followup	Interview Day	none	28.8	0.5
Combined	Interview Day	Estimation *	5.5	0.1
Combined	Interview Day	Sample Cluster	5.5	0.1
Combined	Interview Day	none	5.8	0.1

* Note: The estimation weights most applicable to Interview Day data were the ones generated in the Non-Interview Adjustment operation.

The patterns of proxy results for housing unit data shown in Table 4.1.2 were defined in the same way as above, except that they were created from housing-unit level data. Substitute “housing units” for “persons” in the definitions above.

Table 4.1.2. Percentages of Housing Units Reported by Proxy for Alternative Interview Operations, Reference Days, and Weights Applied

Operation	Reference Day	Weights Applied	Percent of Proxy	s.e.
Original	Census Day	Sample Cluster	7.5	0.1
Original	Census Day	none	7.8	0.1
Followup	Census Day	Sample Cluster	33.5	0.5
Followup	Census Day	none	33.2	0.5
Original	Interview Day	Sample Cluster	4.8	0.1
Original	Interview Day	none	5.0	0.1
Followup	Interview Day	Sample Cluster	30.7	0.5
Followup	Interview Day	none	30.6	0.5

The percentages of proxy response were much greater for followup interviews than for original interviews. In both tables, the differences in similarly weighted percents were all statistically significant. Proxies in the original interview were prominent among followup cases, largely because all proxy nonmatches were sent to followup. The conditions making them proxy in the first interview may also have made them proxy in the followup. Note that analyses of followup interview data were limited to persons with known proxy status. Proxy status was unknown for about 15 percent of the followup cases, due to the relatively low priority of proxy status in the followup data collection.

The differences between Census Day and A.C.E. Interview Day percentages of proxy were small but consistent and statistically significant. Census Day percentages of proxy were higher, likely because those data involved outmovers who, by definition, were harder to find for the interview.

Weighting alternatives had very small effects on the results. Some analyses following used the sample cluster weights rather than the full estimation weights in order to include A.C.E. Interview Day persons. The results should not be viewed an artifact of the weights employed.

Putting followup together with original interview proxy results into a combined percent-of-proxy statistic masked and confounded the separate effects. The original interview effects tended to overwhelm the followup impact in combined rates. Combined interview percents shown in Table 4.1.1 were slightly elevated but very close to corresponding original interview results, but were significantly different from similarly weighted other proxy statistics. Combined interview percentages involved a crude imputation for followup data with unresolved proxy status (essentially using the percent of proxy among resolved cases.) It was not taken into account in estimating standard errors. Combined interview percents were not employed in further analyses.

4.2 What variables were related to the percent of proxy reporting?

This section provides results for relating other characteristics of clusters, households, or persons to percent of proxies in estimation-weighted Census-Day person data for original and for followup interviews. These results help develop an understanding of how proxy reporting interacts with variables that are prominent in coverage measurement. These results do not support conclusions that proxy reporting was caused by or was the cause of another variable's status. But, observed relationships to other variables may lead to speculative interpretations worthy of further analysis.

Results are presented in tables reporting results for each group defined by the selected variable's value level and displaying columns for:

- group name,
- percent of proxies (under the heading "Percent"),
- the rank of that level's percent from lowest to highest ("Rank"),
- a list of the rank numbers of other groups with which a significant difference was found ("Differs from"),
- the stratified jackknife standard error ("s.e."), and
- the weighted percent of persons comprising that level's subgroup ("n(percent)").

The denominator of percents in the "n" column is the weighted total number of persons in the sample (258,547,382 for analyses of original interview data; 12,785,235 for followup interviews.) Percents may not sum precisely to 100 due to rounding error. Criterion t-values (e.g., " $|t| > 1.65$ ") were noted below each table and varied, as described above, with the number of comparisons being made in the family of tests.

Important variables were grouped into two categories:

- variables used in defining post-strata for dual system estimation, and
- other variables relevant to proxy and coverage measurement operations.

4.2.1 *What post-stratification variables were related to the percent of proxy reporting?*

Variables used to form post-strata in dual system estimation (Haines, 2001) were of primary interest. They were analyzed here using the levels as defined for post-stratification. Levels of Metropolitan Statistical Area / Type of Enumeration Area (MSA/TEA), Return Rate Indicator, and Region were used in some but not all post-stratum group definitions. In the following analyses, all P-sample persons were assigned to groups on the basis of MSA/TEA or Region variable values, regardless of whether their post-stratum was differentiated by the variable. In contrast, for the Return Rate Indicator, persons or housing units in poststrata collapsed over census return rate were put into an additional separate analysis group.

Table 4.2.1 shows the percentages of proxy response for the original interview of Census Day persons categorized by the Age/Sex classification of the person. Percentages of proxy were significantly different across all groupings except between 50+ males and females. If we accept that 18-29 year old males and females are more often active and living alone, their higher percent of proxy may have been due to a higher likelihood of no one being at home on successive interview attempts. Similarly, if households containing children or adults aged fifty or older are more likely to have a household member available to respond, a significantly lower proxy rate for

the 0-17 and the 50+ categories would be fitting. Alternatively, if for any reason proxy respondents would be more likely to omit younger children or older adults when providing information about the household, we would see fewer proxies associated with these categories. These alternative explanations would be confounded with each other. Further investigations would be needed to confirm either.

Table 4.2.1. Percent of Original Interview Census Day Person Proxy, by Age and Sex

Age/Sex	Percent	Rank	Differs from	s.e.	n (percent)
0-17 Male & Female	4.3	1	all	0.1	26.2
18-29 Male	7.2	6	all	0.3	7.5
18-29 Female	8.0	7	all	0.4	7.7
30-49 Male	6.2	5	all	0.1	15.2
30-49 Female	5.7	4	all	0.1	16.2
50+ Male	5.0	3	1, 4-7	0.1	12.3
50+ Female	4.8	2	1, 4-7	0.1	15.0

Note: Criterion for levels to differ was $|t| > 2.815$

Table 4.2.2 shows percentages of proxy response for the followup interview for Census Day persons, categorized by Age/Sex. As for the overall data (5.5 percent proxy in the original interview and 32.4 percent proxy in the followup), followup interview results in each age and sex group were substantially higher than original interview results, perhaps because all proxy nonmatches were sent to followup. Again, within this table, we see the same trend of increased proxy response for the males and females aged 18-29 compared to significantly lower proxy response rates for 0-17 and 50+ males and females. When considering the results in Table 4.2.2 relative to those in Table 4.2.1 above, note that there were fewer significant differences among proxy results for Age/Sex categories in the followup interviews than in the original interviews.

Table 4.2.2. Percent of Followup Interview Census Day Person Proxy, by Age and Sex

Age/Sex	Percent	Rank	Differs from	s.e.	n (percent)
0-17 Male & Female	26.2	1	3-7	0.8	27.2
18-29 Male	37.0	6	1-3	1.1	13.3
18-29 Female	37.5	7	1-3	1.2	11.0
30-49 Male	36.9	5	1-3	0.9	15.9
30-49 Female	34.3	4	1, 2	1.0	13.0
50+ Male	31.8	3	1, 5-7	1.2	9.2
50+ Female	28.7	2	4-7	1.1	10.4

Note: Criterion for levels to differ was $|t| > 2.815$

Table 4.2.3 shows percentages of proxy response for the original interview of Census Day persons, categorized by tenure. The percentage of proxy was significantly higher for non-owners than for owners.

Table 4.2.3. Percent of Original Interview Census Day Person Proxy, by Tenure

Tenure	Percent	Rank	Differs from	s.e.	n (percent)
Owner	3.3	1	2	0.1	69.8
Non-owner	10.5	2	1	0.2	30.2

Note: Criterion for levels to differ was $|t| > 1.645$

Table 4.2.4 shows percentages of proxy response for the followup interview of Census Day persons, categorized by tenure. Again, non-owners have a significantly higher percent proxy, just as for the original interview.

Table 4.2.4. Percent of Followup Interview Census Day Person Proxy, by Tenure

Tenure	Percent	Rank	Differs from	s.e.	n (percent)
Owner	20.7	1	2	0.6	52.1
Non-owner	45.1	2	1	0.9	47.9

Note: Criterion for levels to differ was $|t| > 1.645$

Table 4.2.5 shows percentages of proxy response for the original interview of Census Day persons, categorized by Race and Hispanic Origin. Only Non-Hispanic Blacks differed from other groups; they had a significantly higher percent of proxy than any other group.

Table 4.2.5. Percent of Original Interview Census Day Person Proxy, by Race and Hispanic Origin

Race and Hispanic Origin Domain	Percent	Rank	Differs from	s.e.	n (percent)
American Indian on Reservation	4.6	3	7	0.5	0.2
American Indian off Reservation	4.5	2	7	0.6	0.5
Hispanic	4.8	4	7	0.2	12.3
Non-Hispanic Black	6.8	7	all	0.2	11.4
Native Hawaiian or Pacific Islander	4.3	1	7	0.8	0.2
Non-Hispanic Asian	5.0	5	7	0.3	3.4
Non-Hispanic White	5.4	6	7	0.1	72.1

Note: Criterion for levels to differ was $|t| > 2.815$

Table 4.2.6 shows percentages of proxy response for the followup interview of Census Day persons, categorized by Race and Hispanic Origin. Non-Hispanic Blacks did not differ from other groups as in the original interview. Only two groups were statistically significantly different in the followup data; Non-Hispanic Whites had a higher percent of proxy than did Hispanics. They were the largest groups with the smallest variances. They did not have the most extreme percent of proxy results.

Table 4.2.6. Percent of Followup Interview Census Day Person Proxy, by Race and Hispanic Origin

Race and Hispanic Origin Domain	Percent	Rank	Differs from	s.e.	n(percent)
American Indian on Reservation	24.0	1	none	4.1	0.3
American Indian off Reservation	37.3	7	none	4.2	0.7
Hispanic	29.6	3	6	1.2	18.5
Non-Hispanic Black	31.4	4	none	1.1	18.1
Native Hawaiian or Pacific Islander	32.2	5	none	7.5	0.4
Non-Hispanic Asian	28.9	2	none	2.0	3.9
Non-Hispanic White	33.9	6	3	0.8	58.2

Note: Criterion for levels to differ was $|t| > 2.815$

Metropolitan statistical areas (MSAs) delineate cities for statistical purposes. Type of Enumeration Area (TEA) describes the method of data collection adopted for an area. About 82 percent of the sample persons live in Mailout/Mailback (MO/MB) type of enumeration areas, where Census 2000 forms were mailed to their addresses with directions to return responses by

mail. Other TEAs involve Census 2000 staff bringing forms to housing units in the area, usually updating address listings and leaving the forms for mailback, but sometimes listing addresses for the first time or enumerating persons on the spot. Metropolitan Statistical Area size and Type of Enumeration Area were combined into one variable (MSA/TEA) used for post-stratification.

Table 4.2.7 shows percentages of proxy response for the original interview of Census Day persons, categorized by size of the MSA and TEA. One statistically significant difference was found between the non-MO/MB and the Small MSA and Non-MSA MO/MB groups. Alternative explanations include that in areas where mailout and mailback type of enumeration were not adopted, that it might not have been necessary to find proxies, and, on the other hand, that it might have been harder to locate a proxy.

Table 4.2.7. Percent of Original Interview Census Day Person Proxy, by Size of Metropolitan Statistical Area and Type of Enumeration Area

MSA/TEA	Percent	Rank	Differs from	s.e.	n (percent)
Large MSA, MO/MB	5.4	2	none	0.2	30.5
Medium MSA, MO/MB	5.5	3	none	0.2	31.3
Small MSA & Non-MSA, MO/MB	5.8	4	1	0.3	20.2
All Other TEAs, Non-MO/MB	5.0	1	4	0.2	18.1

Note: Criterion for levels to differ was $|t| > 2.386$

Table 4.2.8 shows percentages of proxy response for the followup interview of Census Day persons, categorized by size of the MSA and TEA. The group of All Other TEAs had the lowest percent of proxy as in the analyses of original interview data, but it was statistically significant only against medium MSAs, which also differed from large MSAs.

Table 4.2.8. Percent of Followup Interview Census Day Person Proxy, by Size of Metropolitan Statistical Area and Type of Enumeration Area

MSA/TEA	Percent	Rank	Differs from	s.e.	n (percent)
Large MSA, MO/MB	30.9	2	4	1.0	33.5
Medium MSA, MO/MB	35.7	4	1, 2	1.2	30.2
Small MSA & Non-MSA, MO/MB	32.8	3	none	1.4	18.0
All Other TEAs, Non-MO/MB	29.4	1	4	1.5	18.2

Note: Criterion for levels to differ was $|t| > 2.386$

The tract-level census return rate, a sign of public cooperation, was the proportion of occupied housing units in a census tract that returned a Census 2000 questionnaire. High and low Return Rate Indicator values were assigned for the Non-Hispanic White or “Some other race,” Non-Hispanic Black, and Hispanic domains. Persons in all other Race and Hispanic Origin

Domains were assigned a Return Rate Indicator value of “Not Applicable” since they were not post-stratified by return rate (Haines, 2001.)

Table 4.2.9 shows percentages of proxy response for the original interview of Census Day persons, categorized by tract-level Return Rate Indicator. Areas recognized as having low census return rates also had higher percent of proxy results for the original interview.

Table 4.2.9. Percent of Original Interview Census Day Person Proxy, by Return Rate Indicator

Return Rate Indicator	Percent	Rank	Differs from	s.e.	n (percent)
High	5.2	2	3	0.1	72.3
Low	6.5	3	all	0.2	23.5
Not Applicable	4.9	1	3	0.3	4.3

Note: Criterion for levels to differ was $|t| > 2.121$

Table 4.2.10 shows percentages of proxy response for the followup interview of Census Day persons, categorized by Return Rate Indicator. No statistically significant differences were evident for the followup interview. Perhaps the followup data were already so loaded with persons and households needing proxy help, the impact of census cooperation was diminished.

Table 4.2.10. Percent of Followup Interview Census Day Person Proxy, by Return Rate Indicator

Return Rate Indicator	Percent	Rank	Differs from	s.e.	n (percent)
High	32.3	2	none	0.7	64.2
Low	33.1	3	none	1.1	30.5
Not Applicable	29.9	1	none	1.8	5.3

Note: Criterion for levels to differ was $|t| > 2.121$

Table 4.2.11 shows percentages of proxy response for the original interview of Census Day persons, categorized by region of the United States. The South and West regions had statistically higher percent of proxy results than the Northeast and Midwest.

Table 4.2.11. Percent of Original Interview Census Day Person Proxy, by Region of the United States

Region	Percent	Rank	Differs from	s.e.	n (percent)
Northeast	5.0	2	3, 4	0.2	19.0
Midwest	4.5	1	3, 4	0.2	22.9
South	6.1	4	1, 2	0.2	35.4
West	5.9	3	1, 2	0.2	22.8

Note: Criterion for levels to differ was $|t| > 2.386$

Table 4.2.12 shows percentages of proxy response for the followup interview of Census Day persons, categorized by region of the United States. The South and West regions had statistically higher percent of proxy results than the Northeast, but the difference with the Midwest was not statistically significant in the followup data as it was in the original interview data.

Table 4.2.12. Percent of Followup Interview Census Day Person Proxy, by Region of the United States

Region	Percent	Rank	Differs from	s.e.	n (percent)
Northeast	28.8	1	3, 4	1.2	18.9
Midwest	31.7	2	none	1.2	17.3
South	33.3	3	1	1.0	39.3
West	34.3	4	1	1.4	24.6

Note: Criterion for levels to differ was $|t| > 2.386$

4.2.2 What other variables were related to the percent of proxy reporting?

Other operational or characteristic variables were analyzed, as described below, including: imputation of characteristics, type of structure at basic street address, involvement in the Targeted Extended Search (TES) operation at person and cluster levels, match status, mover status, household size, relationship to reference person, and Accuracy and Coverage Evaluation sample exclusion status.

Tables 4.2.13 and 4.2.14 show percentages of proxy response for Census Day persons, categorized by imputation of characteristics. Age, sex, race, Hispanic origin, and tenure were sometimes imputed for the A.C.E. Persons for whom major characteristics (tenure, age, sex, race, and Hispanic origin) were imputed had a much higher percentage of proxy responses than persons with no imputed characteristics in both original (21.6 and 4.6 percents, respectively) and followup (41.0 and 30.7 percents, respectively) interviews. Complete data were associated with

non-proxy respondents, who presumably were more familiar with characteristics of the household members.

Table 4.2.13. Percent of Original Interview Census Day Person Proxy, by Imputation of Characteristics

Imputed or Not	Percent	Rank	Differs from	s.e.	n (percent)
Not Imputed	4.6	1	2	0.1	94.7
Imputed	21.6	2	1	0.7	5.3

Note: Criterion for levels to differ was $|t| > 1.645$

Table 4.2.14. Percent of Followup Interview Census Day Person Proxy, by Imputation of Characteristics

Imputed or Not	Percent	Rank	Differs from	s.e.	n (percent)
Not Imputed	30.7	1	2	0.6	83.0
Imputed	41.0	2	1	1.3	17.0

Note: Criterion for levels to differ was $|t| > 1.645$

Tables 4.2.15 and 4.2.16 show percentages of proxy response for Census Day persons, categorized by type of structure at the basic street address. Type of structure at a basic street address distinguishes single-family dwellings from multi-unit buildings, mobile homes and living quarters in a special place as well as unclassified structures. If the small number of living quarters in a special place and unclassified structures are ignored because they are so variable, single-family dwellings had the smallest percent of proxies and multi-units the largest, while mobile homes were between them in both original and followup data collections.

Table 4.2.15. Percent of Original Interview Census Day Person Proxy, by Type of Structure at Basic Street Address

Structure	Percent	Rank	Differs from	s.e.	n (percent)
Single-Family Dwelling	3.7	1	all	0.1	75.7
Multi-Unit	12.0	3	1, 2	0.3	18.9
Mobile Home	6.9	2	1, 3	0.3	5.2
Living Quarters in a Special Place and Unclassified	13.5	4	1	2.9	0.2

Note: Criterion for levels to differ was $|t| > 2.386$

Table 4.2.16. Percent of Followup Interview Census Day Person Proxy, by Type of Structure at Basic Street Address

Structure	Percent	Rank	Differs from	s.e.	n (percent)
Single-Family Dwelling	25.0	1	3, 4	0.6	59.9
Multi-Unit	46.1	4	all	1.2	31.9
Mobile Home	33.5	3	1, 4	2.0	7.8
Living Quarters in a Special Place and Unclassified	28.9	2	4	6.1	0.4

Note: Criterion for levels to differ was $|t| > 2.386$

Tables 4.2.17 and 4.2.18 show percentages of proxy response for Census Day persons, categorized by TES Status, whether or not the person was part of the Targeted Extended Search operation designed to measure and account for geocoding error. In the original interview, TES persons had a statistically significantly higher percentage of proxy reporting than did persons not in the TES operation. There was no such difference in the followup data. Since TES persons were identified partly on nonmatch status, these findings may have been a mere reflection of the match status findings below.

Table 4.2.17. Percent of Original Interview Census Day Person Proxy, by TES Status

TES Status	Percent	Rank	Differs from	s.e.	n (percent)
TES Person	8.4	2	1	0.5	6.0
Not TES Person	5.3	1	2	0.1	94.0

Note: Criterion for levels to differ was $|t| > 1.645$

Table 4.2.18. Percent of Followup Interview Census Day Person Proxy, by TES Status

TES Status	Percent	Rank	Differs from	s.e.	n (percent)
TES Person	32.7	2	none	2.2	14.6
Not TES Person	32.4	1	none	0.6	85.4

Note: Criterion for levels to differ was $|t| > 1.645$

Tables 4.2.19 and 4.2.20 show percentages of proxy response for Census Day persons, categorized by whether they lived in a cluster involved in the TES or address relisting operations conducted when the quality of the cluster's original address listing was called into question. In the original interview, persons in certainty TES clusters, selected without fail just because that cluster had the highest incidence of nonmatched persons, had a statistically significantly higher percentage of proxy reporting than did persons in clusters that were sampled or not in the TES operation at all. There were no differences in the followup data.

Table 4.2.19. Percent of Original Interview Census Day Person Proxy, by Targeted Extended Search and Relist Cluster Status

Cluster Status	Percent	Rank	Differs from	s.e.	n (percent)
Relisted Cluster	12.8	4	none	3.8	0.2
Certainty TES Cluster	6.3	3	1, 2	0.3	14.7
Sample TES Cluster	5.4	2	3	0.3	12.0
Out of TES Sample or Scope	5.3	1	3	0.1	73.0

Note: Criterion for levels to differ was $|t| > 2.386$

Table 4.2.20. Percent of Followup Interview Census Day Person Proxy, by Targeted Extended Search and Relist Cluster Status

Cluster Status	Percent	Rank	Differs from	s.e.	n (percent)
Relisted Cluster	36.8	4	none	9.1	0.6
Certainty TES Cluster	33.7	3	none	1.6	18.7
Sample TES Cluster	29.6	1	none	2.0	13.0
Out of TES Sample or Scope	32.6	2	none	0.7	67.7

Note: Criterion for levels to differ was $|t| > 2.386$

Tables 4.2.21 and 4.2.22 show percentages of proxy response for the original and followup interview of Census Day persons, categorized by match status. Matched persons had statistically significantly lower percentages of proxy response than nonmatched persons in the original interview, but the relationship was reversed for the followup. Note the low percentage of matches involved in the followup relative to that in the original interview. Matches generally do not need followup. Matches among followup cases were mostly converted from unresolved or nonmatched persons in the original interview.

Both matches and nonmatches had lower percentages of proxy respondents than those for whom match status had to be imputed, although in the followup operation that difference was statistically significant for nonmatches only. These findings seem to suggest that resolved match status may have been more likely found, especially in the original interview, without need of proxy or, in another view, that proxies may have had too little information, resulting in more nonmatches or unresolved statuses. The contradiction in those two interpretations underscores the need for caution in explaining either proxy status or match status as the cause of the other.

Table 4.2.21. Percent of Original Interview Census Day Person Proxy, by Match Status

Match Status	Percent	Rank	Differs from	s.e.	n (percent)
Match	4.3	1	all	0.1	91.0
Imputed Status	55.1	3	all	1.2	1.0
Nonmatch	12.6	2	all	0.3	8.0

Note: Criterion for levels to differ was $|t| > 2.121$

Table 4.2.22. Percent of Followup Interview Census Day Person Proxy, by Match Status

Match Status	Percent	Rank	Differs from	s.e.	n (percent)
Match	38.6	2	1	1.2	19.4
Imputed Status	50.1	3	1	6.5	0.3
Nonmatch	30.8	1	all	0.6	80.3

Note: Criterion for levels to differ was $|t| > 2.121$

Tables 4.2.23 and 4.2.24 show percentages of proxy response for the original and followup interviews of Census Day persons, categorized by mover status. In both operations, outmovers had much larger percent of proxy values. That would be mostly due to whole-household outmovers, who were captured on the CAPI instrument's outmover path and thereby assigned proxy status.

Table 4.2.23. Percent of Original Interview Census Day Person Proxy, by Mover Status

Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Nonmover	3.0	1	2	0.1	96.6
Outmover	74.8	2	1	0.6	3.4

Note: Criterion for levels to differ was $|t| > 1.645$

Table 4.2.24. Percent of Followup Interview Census Day Person Proxy, by Mover Status

Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Nonmover	26.2	1	2	0.6	85.8
Outmover	70.3	2	1	1.2	14.2

Note: Criterion for levels to differ was $|t| > 1.645$

Tables 4.2.25 and 4.2.26 show percentages of proxy response for Census Day persons, categorized by household size. In both original and followup interviews, there was a strong relationship between the number of persons in the household and proxy status – the fewer the

number of household members the greater the percentage of proxy response. More household members might make it more likely that interviewers will find a household respondent. An alternate explanation might be that proxy respondents reported fewer household members.

Table 4.2.25. Percent of Original Interview Census Day Person Proxy, by Household Size

Household Size	Percent	Rank	Differs from	s.e.	n (percent)
One person	13.5	3	all	0.2	10.5
2-6 persons	4.6	2	all	0.1	84.9
7 or more persons	2.5	1	all	0.3	4.7

Note: Criterion for levels to differ was $|t| > 2.121$

Table 4.2.26. Percent of Followup Interview Census Day Person Proxy, by Household Size

Household Size	Percent	Rank	Differs from	s.e.	n (percent)
One person	57.9	3	all	1.1	12.0
2-6 persons	30.6	2	all	0.7	77.6
7 or more persons	16.4	1	all	1.4	10.4

Note: Criterion for levels to differ was $|t| > 2.121$

Tables 4.2.27 and 4.2.28 show percentages of proxy response for Census Day persons, categorized by kinship to the reference person. In both original and followup interviews, the reference person who had no other household members had the highest percent proxy. This was essentially the same finding as for household size of one person above. Within households of larger size, multi-generation families yielded the lowest percentage of proxies. One explanation might be that someone was more likely to be home to respond to the interview, but it could also be that some proxies did not know of, or think of, or want to report on everyone. The other kinship categories were not consistent from original to followup interview, but may reflect varying degrees of the availability of a household member to be a respondent.

Table 4.2.27. Percent of Original Interview Census Day Person Proxy, by Kinship to Reference Person

Kinship to Reference Person	Percent	Rank	Differs from	s.e.	n (percent)
Reference Person, Alone	13.5	5	all	0.2	10.1
Reference Person, Not Alone	5.1	3	all	0.1	28.3
Spouse	4.1	2	all	0.1	20.2
Parent / Child	3.8	1	all	0.1	30.4
Other relatives and nonrelatives	6.2	4	all	0.3	10.9

Note: Criterion for levels to differ was $|t| > 2.568$

Table 4.2.28. Percent of Followup Interview Census Day Person Proxy, by Kinship to Reference Person

Kinship to Reference Person	Percent	Rank	Differs from	s.e.	n (percent)
Reference Person, Alone	58.0	5	all	1.1	11.4
Reference Person, Not Alone	37.6	4	all	0.9	20.3
Spouse	29.2	3	1, 4, 5	1.0	12.6
Parent / Child	24.4	1	all	0.8	29.2
Other relatives and nonrelatives	27.8	2	1, 4, 5	0.9	26.5

Note: Criterion for levels to differ was $|t| > 2.568$

Tables 4.2.29 and 4.2.30 show percentages of proxy response for Census Day persons, categorized by A.C.E. P-sample exclusion status. A small number of persons were determined to be nonresidents or fictitious and thus deleted from the sample. The deletes were given an estimation weight equal to zero, so a weight (PWGHT) from a prior stage of processing was used in place of the final estimation weight. While the cases removed had higher percent of proxy values in original interview data, the comparison result was reversed in followup, statistically significant in each case. One explanation might be that a proxy respondent was more likely to falsely enumerate a person in the original interview, and it took a non-proxy to ascertain that they should be removed in the followup interview.

Table 4.2.29. Percent of Original Interview Census Day Person Proxy, by A.C.E. Sample Exclusion Status

A.C.E. Sample Exclusion	Percent	Rank	Differs from	s.e.	n (percent)
Coded to Remove	8.4	2	1	0.6	0.5
Not Removed	5.5	1	2	0.1	99.5

Notes: Criterion for levels to differ was $|t| > 1.645$

Sample cluster weight used instead of final estimation weight.

Weighted sample size is 265,042,046.

Table 4.2.30. Percent of Followup Interview Census Day Person Proxy, by A.C.E. Sample Exclusion Status

A.C.E. Sample Exclusion	Percent	Rank	Differs from	s.e.	n (percent)
Coded to Remove	15.4	1	2	0.9	8.3
Not Removed	33.5	2	1	0.6	91.7

Notes: Criterion for levels to differ was $|t| > 1.645$

Sample cluster weight used instead of final estimation weight.

Weighted sample size is 15,180,204.

4.3 How did the percent of proxy enumeration relate to consistency between matched Census 2000 and A.C.E. data?

The percent of proxy reporting for people who had the same response values to Census 2000 and A.C.E. items was compared to the percent for those with inconsistent answers. The tables are laid out just as in the previous section. One comparison group in each table consisted of persons with consistent responses on a selected item; the other group was formed from those whose responses differed. Data were restricted to the 222,922,906 estimation-weighted matched Census Day persons for whom we had both Census 2000 and A.C.E. responses.

Tables 4.3.1 to 4.3.5 present the percent of proxy comparisons for consistency of Census Day person responses to age and sex categories, tenure, race and Hispanic origin domain, type of structure at basic street address, and household size categories. In each of these tests, when the Census 2000 and A.C.E. responses did not agree, the percentage of proxy responses was higher and the difference was statistically significant. One explanation might be that proxies are more likely to give inconsistent data; another might be that conditions that lead to inconsistent data also lead to the need for proxies.

Table 4.3.1. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Age and Sex Categories

Age/Sex Consistency	Percent	Rank	Differs from	s.e.	n (percent)
Consistent	3.6	1	2	0.1	95.0
Different	13.5	2	1	0.3	5.0

Note: Criterion for levels to differ was $|t| > 1.645$

Table 4.3.2. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Tenure Categories

Tenure Consistency	Percent	Rank	Differs from	s.e.	n (percent)
Consistent	4.0	1	2	0.1	95.8
Different	6.6	2	1	0.3	4.2

Note: Criterion for levels to differ was $|t| > 1.645$

Table 4.3.3. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Race and Hispanic Origin Domain Categories

Race and Hispanic Origin Domain Consistency	Percent	Rank	Differs from	s.e.	n (percent)
Consistent	3.9	1	2	0.1	96.4
Different	7.9	2	1	0.3	3.6

Note: Criterion for levels to differ was $|t| > 1.645$

Table 4.3.4. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Type of Structure at Basic Street Address Categories

Type of Structure Consistency	Percent	Rank	Differs from	s.e.	n (percent)
Consistent	4.0	1	2	0.1	97.4
Different	6.2	2	1	0.4	2.6

Notes: Criterion for levels to differ was $|t| > 1.645$

Since the Census 2000 categorized all records as single-family or multi-unit, A.C.E. records in other groups were not included in the analysis; estimation-weighted sample size was 212,161,939.

Table 4.3.5. Percent of Proxy among Matched Persons, by Consistency of Census and A.C.E. Household Size Categories

Household Size Consistency	Percent	Rank	Differs from	s.e.	n (percent)
Consistent	3.9	1	2	0.1	96.7
Different	10.1	2	1	0.4	3.3

Note: Criterion for levels to differ was $|t| > 1.645$

4.4 Did proxy response on the original A.C.E. interview relate to proxy response in the followup interview?

This section presents results on the proxy status in the original interview combined with the proxy status in the followup. The tables, except for the full sample summary of those proxy status combinations, display the proxy status combination counts within groups defined on variables as in section 4.2 above. Data were restricted to the 12,785,235 estimation-weighted Census Day persons with known proxy status in both the original and followup operations.

Table 4.4.1 shows there was a relationship in proxy response from one operation to the other, as indicated by a statistically significant result in a Chi-square test of independence. Looking at specific cells shows most of the cases (58.5 percent) had no proxy at all. Among the persons who did not have proxy response in the original interview, a small portion (about 21 percent) had a proxy report in followup. In contrast, among persons who had a proxy response in the original interview, a large portion (about 65 percent) had proxy response in the followup.

Table 4.4.1. Frequencies and Percents of Persons with Proxy Report: Original Interview by Followup Interview

	No FU Proxy	FU Proxy	Total
No Original Proxy	7,485,747 58.6%	2,029,567 15.9%	9,515,314 74.4%
Original Proxy	1,155,646 9.0%	2,114,275 16.5%	3,269,921 25.6%
Total	8,641,393 67.6%	4,143,842 32.4%	12,785,235 100.0%

Note: Chi-square probability of independence was < 0.0001

In Tables 4.4.2 to 4.4.4, the four cells of table 4.4.1 are rearranged into column headings. Each row is a group defined from values of Tenure, Race and Hispanic Origin Domain, and Age/Sex categories. To better identify when patterns within groups are unlike the overall pattern as in Table 4.1.1, the percents in the following tables are computed within groups. The final column (“n (percent)”), represents the portion of all followup cases belonging to that row.

In Table 4.4.2, we see the effects are not the same for owners as for non-owners. Non-owners in the original interview appeared more likely to have had followup proxy or proxy on both interviews, but not just original interview proxy.

Table 4.4.2. Percent of Persons By Proxy Status Combination within Tenure Groups

Tenure	No Proxy	Original Proxy	FU Proxy	Both Proxy	n (percent)
Owner	69.6	9.7	11.4	9.3	52.1
Non-owner	46.6	8.3	20.7	24.4	47.9
Total	58.5	9.0	15.9	16.5	100.0

Note: Chi-square probability of independence was < 0.0001

Again in Table 4.4.3, the test of independence over the whole table shows that, from one level of Race and Hispanic Origin Domain to another, the distribution across proxy status combinations were not the same. In all groups, as for the overall total, the No Proxy combination had the largest percent and Original Proxy had the smallest percent of persons. FU Proxy and Both Proxy percents were nearly the same in the total, but individual race groups generally had greater disparities between the two, with only Non-Hispanic Whites yielding more Both Proxy persons.

Table 4.4.3. Percent of Persons By Proxy Status Combination within Race and Hispanic Origin Domain Groups

Race and Hispanic Origin Domain	No Proxy	Original Proxy	FU Proxy	Both Proxy	n (percent)
American Indian on Reservation	73.2	2.9	18.9	5.0	0.3
American Indian off Reservation	59.8	2.9	24.9	12.4	0.7
Hispanic	65.9	4.5	18.8	10.8	18.5
Non-Hispanic Black	59.7	8.9	17.0	14.4	18.1
Native Hawaiian or Pacific Islander	64.8	3.0	25.8	6.4	0.4
Non-Hispanic Asian	64.6	6.5	16.7	12.1	3.9
Non-Hispanic White	55.3	10.8	14.4	19.5	58.2
Total	58.5	9.0	15.9	16.5	100.0

Note: Chi-square probability of independence was < 0.0001

In Table 4.4.4, the test of independence over the whole table shows that, from one age and sex category to another, the distribution across proxy status combinations were not the same. Children's relatively low proxy rate in original interview shows up in the high no-proxy percentage. Men and women aged 30 or more had relatively more original-interview-only proxy reports, while younger groups tended to have somewhat more followup-only proxies.

Table 4.4.4. Percent of Persons By Proxy Status Combination within Age and Sex Groups

Age / Sex	No Proxy	Original Proxy	FU Proxy	Both Proxy	n (percent)
0 - 17 Male and Female	68.3	5.5	16.8	9.4	27.2
18 - 29 Male	58.1	4.9	19.4	17.6	13.3
18 - 29 Female	55.6	6.9	16.6	20.9	11.0
30 - 49 Male	52.5	10.5	16.2	20.8	15.9
30 - 49 Female	52.3	13.4	14.1	20.2	13.0
50+ Male	54.0	14.1	13.4	18.4	9.2
50+ Female	57.8	13.5	12.2	16.5	10.4
Total	58.5	9.0	15.9	16.5	100.0

Note: Chi-square probability of independence was < 0.0001

4.5 Did the percent of Item Non-response in the original A.C.E. interview relate to the source of proxy response?

This section presents results on the percent of item non-response, including refusals and item-missing as well as “Don’t Know” responses, to specific items in the original interview within groups defined by the type of proxy respondent (Landlord, Neighbor, Postal Worker, and Observation or Other), plus a group for non-proxy respondents. See the appendix for analyses of all proxy versus non-proxy percent of item non-response. The percent of item non-response is one measure, even if not a complete one, of the quality of a response. “Don’t Know,” blank, and missing values for a question may represent unwillingness to answer a particular question, some degree of uncertainty about the answer, or outright lack of knowledge, so those values are pooled with refusals to represent this kind of inadequate response. These comparisons show how well different types of respondents were able to answer specific items. Except as noted for the last two tables of this section, computations were based on the 271,449,765 cluster-weighted complete interviews with both Census Day and A.C.E. Interview Day persons in the original interview. In the situation where an address had both in-mover and out-mover households and one was not a complete interview, only the household with complete data was included.

For the variety of items analyzed in Tables 4.5.1 to 4.5.8, the non-proxy respondent group generally had the smallest percentage of item non-response; household members were the most knowledgeable and willing respondents. In the analyses reported in the appendix, non-proxy respondents always yielded statistically significantly smaller percentages of item non-response than proxy respondents as a group. Postal workers often yielded the highest percentage of item non-response, even if not always with statistically significant differences. In general, they appeared to be the most unsure or cautious in responding. Landlords, with one exception, also often yielded a high or the highest percentage of non-response. Results from neighbors tended to be in the middle of the range. The “interviewer observation and other” source of response was a mixed group, hard to characterize or interpret simply; that group’s results were relatively less consistent. See the descriptions below regarding specific items’ percent of item non-response.

In Table 4.5.1, landlords were about as certain as household respondents concerning whether the residence is owned or not. Having a landlord relationship to the sample household should explain knowing renter status. That was not generally so for more personal characteristics.

Table 4.5.1. Percent of Item Non-response to the Tenure Item, by Type of Proxy Respondent

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.5	2	3, 4, 5	<0.1	94.5
Landlord	0.3	1	3, 4, 5	0.1	1.8
Neighbor	5.3	4	all	0.4	1.7
Postal Worker	23.5	5	all	4.0	0.1
Observation or Other	1.8	3	all	0.2	1.9

Note: Criterion for levels to differ was $|t| > 2.568$

In Table 4.5.2 the percentages of item non-response to age were statistically significantly lower for the non-proxy group and higher for postal workers than all other groups. Other differences were not statistically significant.

Table 4.5.2. Percent of Item Non-response to the Age Item, by Type of Proxy Respondent

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	2.1	1	all	<0.1	94.5
Landlord	7.6	4	1, 5	0.6	1.8
Neighbor	6.9	2	1, 5	0.4	1.7
Postal Worker	14.5	5	all	2.3	0.1
Observation or Other	7.4	3	1, 5	0.5	1.9

Note: Criterion for levels to differ was $|t| > 2.568$

In Table 4.5.3 the percentages of item non-response regarding gender were all relatively small, perhaps reflecting very little uncertainty among those who know the persons being reported. Landlords yielded the highest percent of item non-response and was significantly statistically different from neighbors' percentages as well as non-proxies.

Table 4.5.3. Percent of Item Non-response to the Gender Item, by Type of Proxy Respondent

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.3	1	all	<0.1	94.5
Landlord	2.4	5	1, 2	0.3	1.8
Neighbor	1.2	2	1, 5	0.2	1.7
Postal Worker	2.1	4	1	0.7	0.1
Observation or Other	1.8	3	1	0.2	1.9

Note: Criterion for levels to differ was $|t| > 2.568$

In Tables 4.5.4 and 4.5.5, the landlords' percentage of item non-response regarding race and Hispanic origin were the highest and were statistically different, except for postal workers, who had a large variance in their estimate. Landlords might have been reticent to report what is sometimes thought to be sensitive information or they might have been uncertain about all the race and ethnicity categories that apply.

Table 4.5.4. Percent of Item Non-response to the Race Item, by Type of Proxy Respondent

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	3.7	1	all	0.1	94.5
Landlord	18.8	5	1, 2, 4	1.6	1.8
Neighbor	7.8	2	1, 4, 5	0.5	1.7
Postal Worker	13.5	3	1	3.3	0.1
Observation or Other	14.0	4	1, 2, 5	1.0	1.9

Note: Criterion for levels to differ was $|t| > 2.568$

Table 4.5.5. Percent of Item Non-response to the Hispanic Origin Item, by Type of Proxy Respondent

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.7	1	all	<0.1	94.5
Landlord	15.5	5	1, 2, 3	1.4	1.8
Neighbor	4.2	2	all	0.4	1.7
Postal Worker	11.3	4	1, 2	2.3	0.1
Observation or Other	9.8	3	1, 2, 5	0.9	1.9

Note: Criterion for levels to differ was $|t| > 2.568$

Tables 4.5.6 to 4.5.8 follow the pattern described at the beginning of this section -- the lowest percentages of item non-response came from non-proxies, the highest from postal workers, while neighbors and landlords share the middle ground. Landlords did give statistically significant higher non-response, relative to neighbors, on the kinship and special place items, but not on the other residence flag. We might speculate that landlords exhibited an enhanced degree of caution or uncertainty about personal questions.

Table 4.5.6. Percent of Item Non-response to the Kinship to Reference Person Item, by Type of Proxy Respondent

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.2	1	all	<0.1	94.5
Landlord	2.5	4	1, 2, 3	0.3	1.8
Neighbor	1.1	2	1, 4, 5	0.2	1.7
Postal Worker	5.5	5	1, 2, 3	1.3	0.1
Observation or Other	1.3	3	1, 4, 5	0.2	1.9

Note: Criterion for levels to differ was $|t| > 2.568$

Additional person records were removed from the analyses of other residence and special place items because the questions were not asked about these persons. The reduction in sample size is noted below each table.

Table 4.5.7. Percent of Item Non-response to the Other Residence Flag Item, by Type of Proxy Respondent

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.6	1	all	<0.1	94.5
Landlord	13.2	4	1, 5	0.9	1.7
Neighbor	12.1	3	1, 5	0.6	1.8
Postal Worker	30.1	5	all	4.1	0.1
Observation or Other	11.2	2	1, 5	0.7	1.9

Note: Criterion for levels to differ was $|t| > 2.568$
 Weighted sample size is 261,506,442.

Table 4.5.8. Percent of Item Non-response to the Special Place Flag, by Type of Proxy Respondent

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.5	1	all	<0.1	94.5
Landlord	10.2	4	1, 2, 5	0.9	1.7
Neighbor	6.9	2	1, 4, 5	0.5	1.8
Postal Worker	23.5	5	all	3.9	0.1
Observation or Other	8.2	3	1, 5	0.7	1.9

Note: Criterion for levels to differ was $|t| > 2.568$
 Weighted sample size is 261,530,078.

4.6 Did the percent of Item Non-response in the original A.C.E. interview relate to the mover status?

This section presents results comparing groups of proxy responses defined by the mover status of the person. Persons in households where everyone moved to a different housing unit after Census Day but before A.C.E. Interview Day are whole-household outmovers. Those in households where everyone moved into the sample housing unit since the Census Day enumeration are whole-household inmovers. Those in households where at least one person was there for both Census Day and A.C.E. Interview Day are classified in one of three ways:

- Nonmovers,
- Inmovers among nonmovers, or
- Outmovers among nonmovers.

In the analyses below, non-proxies of all mover statuses are pooled into one group to put all the relatively small proxy mover groups into context.

Comparisons were based on each group's percent of item non-response to specific items in the

original interview. These comparisons were designed to show whether a respondent might feel more unwillingness, uncertainty, or lack of knowledge about answers for some category of persons who moved in or out of the address around the time of the Census 2000. Computations were based on the 271,449,765 cluster-weighted complete interviews with both Census Day and A.C.E. Interview Day persons in the original interview, as in the previous section.

In Tables 4.6.1 to 4.6.8, non-proxies had the lowest item non-response, as in most of the proxy-type analyses of the previous section. In more than half of those tests, the difference was statistically significant.

Comparing proxy mover groups across the tables below, Inmovers among Nonmovers and Outmovers among Nonmovers seldom had statistically significant differences with other proxy mover groups, except in race and Hispanic origin analyses. Nonmovers, Whole-household Inmovers, and Whole-household Outmovers occasionally had statistically significant differences, but no strong overall trend was evident.

One way to view these analyses is to group items. Sex, kinship, other residence flag, and special place flag had no differences among any mover groups for proxies; tenure and age had few. Race and Hispanic origin had many statistically significant differences.

Table 4.6.1 shows that whole-household mover (both inmover and outmover) households had tenure reported with less item non-response than nonmovers' households.

Table 4.6.1. Percent of Item Non-response to the Tenure Item, by Mover Status

Proxy Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Whole-household Inmovers	0.9	2	5	0.3	0.2
Inmover among Nonmovers	5.3	6	none	3.3	<0.1
Nonmovers	4.0	5	1, 2, 3	0.3	2.8
Outmover among Nonmovers	3.2	4	none	1.4	<0.1
Whole-household Outmovers	1.5	3	1, 5	0.2	2.4
Non-Proxy	0.5	1	3, 5	<0.1	94.5

Note: Criterion for levels to differ was $|t| > 2.705$

The main finding in Table 4.6.2 is that whole-household outmovers had a statistically significant lower percent of item non-response for age than both nonmovers and whole-household inmovers.

Table 4.6.2. Percent of Item Non-response to the Age Item, by Mover Status

Proxy Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Whole-household Inmovers	11.8	5	1, 2	1.4	0.2
Inmover among Nonmovers	8.6	3	none	3.1	<0.1
Nonmovers	8.9	4	1, 2	0.5	2.8
Outmover among Nonmovers	12.7	6	1	2.8	<0.1
Whole-household Outmovers	5.2	2	1, 4, 5	0.3	2.4
Non-Proxy	2.1	1	2, 4, 5, 6	<0.1	94.5

Note: Criterion for levels to differ was $|t| > 2.705$

Overall the only statistically significant differences in Table 4.6.3 were comparisons to non-proxies; proxy mover groups had small and similar percents of item non-response for gender.

Table 4.6.3. Percent of Item Non-response to the Gender Item, by Mover Status

Proxy Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Whole-household Inmovers	2.2	5	1	0.6	0.2
Inmover among Nonmovers	1.6	3	none	1.5	<0.1
Nonmovers	1.5	2	1	0.2	2.8
Outmover among Nonmovers	2.3	6	none	1.0	<0.1
Whole-household Outmovers	2.1	4	1	0.2	2.4
Non-Proxy	0.3	1	2, 4, 5	<0.1	94.5

Note: Criterion for levels to differ was $|t| > 2.705$

The strongest effects among mover status groups and highest percents of item non-response were found in how proxies responded to race and Hispanic origin items. Tables 4.6.4 and 4.6.5 show, for both analyses of these items, that proxies reporting for Outmovers among Nonmovers gave a statistically higher percent of item non-response than every other group. Other whole-household movers, both inmovers and outmovers, followed with the next highest percentages, although not always with statistical significance relative to Nonmovers or Inmovers among Nonmovers. The high rate of non-response on these two items for Outmovers among Nonmovers, who once were in the households and otherwise receive fairly complete item response, suggests that the item might be more difficult for a proxy. If we accept that the proxy is as likely to know this answer as well as other items asked, it seems that willingness is the issue.

Table 4.6.4. Percent of Item Non-response to the Race Item, by Mover Status

Proxy Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Whole-household Inmovers	26.7	5	1, 2, 4, 6	2.2	0.2
Inmover among Nonmovers	14.5	3	6	5.0	<0.1
Nonmovers	10.7	2	1, 4, 5, 6	0.7	2.8
Outmover among Nonmovers	43.2	6	all	5.4	<0.1
Whole-household Outmovers	15.3	4	1, 2, 5, 6	1.1	2.4
Non-Proxy	3.7	1	2, 4, 5, 6	0.1	94.5

Note: Criterion for levels to differ was $|t| > 2.705$

Table 4.6.5. Percent of Item Non-response to the Hispanic Origin Item, by Mover Status

Proxy Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Whole-household Inmovers	11.8	5	1, 2, 6	1.8	0.2
Inmover among Nonmovers	3.4	2	4, 5, 6	2.0	<0.1
Nonmovers	8.5	3	1, 6	0.6	2.8
Outmover among Nonmovers	50.9	6	all	5.3	<0.1
Whole-household Outmovers	10.7	4	1, 2, 6	1.0	2.4
Non-Proxy	0.7	1	3, 4, 5, 6	<0.1	94.5

Note: Criterion for levels to differ was $|t| > 2.705$

Overall the only statistically significant differences in Tables 4.6.6 to 4.6.8 were comparisons to non-proxies. Proxy mover groups also generally had relatively small percents of item non-response for the question concerning kinship. The other residence and special place flag had more and quite variable item non-response.

Table 4.6.6. Percent of Item Non-response to the Kinship Item, by Mover Status

Proxy Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Whole-household Inmovers	1.4	2	none	0.5	0.2
Inmover among Nonmovers	1.5	3	none	1.5	<0.1
Nonmovers	1.5	4	1	0.2	2.8
Outmover among Nonmovers	4.7	6	1	1.5	<0.1
Whole-household Outmovers	1.9	5	1	0.2	2.4
Non-Proxy	0.2	1	4, 5, 6	<0.1	94.5

Note: Criterion for levels to differ was $|t| > 2.705$

Additional person records were removed from the analyses of other residence and special place items because the questions were not asked about these persons. The reduction in sample size is noted below each table.

Table 4.6.7. Percent of Item Non-response to the Other Residence Item, by Mover Status

Proxy Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Whole-household Inmovers	13.5	5	none	8.5	<0.1
Inmover among Nonmovers	11.4	2	none	4.2	<0.1
Nonmovers	12.3	3	1	0.6	2.9
Outmover among Nonmovers	16.4	6	1	4.7	<0.1
Whole-household Outmovers	12.6	4	1	0.6	2.5
Non-Proxy	0.6	1	3, 4, 6	<0.1	94.5

Note: Criterion for levels to differ was $|t| > 2.705$

Weighted sample size is 261,506,442.

Table 4.6.8. Percent of Item Non-response to the Special Place Flag Item, by Mover Status

Proxy Mover Status	Percent	Rank	Differs from	s.e.	n (percent)
Whole-household Inmovers	15.2	6	none	8.2	<0.1
Inmover among Nonmovers	8.6	4	none	3.3	<0.1
Nonmovers	8.4	3	1	0.5	2.9
Outmover among Nonmovers	5.6	2	none	2.0	<0.1
Whole-household Outmovers	9.0	5	1	0.5	2.5
Non-Proxy	0.5	1	3, 5	<0.1	94.5

Note: Criterion for levels to differ was $|t| > 2.705$

Weighted sample size is 261,530,078.

In general, item non-response from one proxy mover group was more pronounced relative to another for only certain items. Race and Hispanic origin questions stood out as worthy of analysis in this regard. From item to item there was no general pattern for any particular mover group to stand out with higher percents of item non-response.

5. CONCLUSIONS AND RECOMMENDATIONS

A few conclusions were drawn from these analyses:

- The percent of proxy responses was smaller for the original A.C.E. interview than for followup and for A.C.E. Interview Day residents than for Census Day residents.
- The overall percentage of proxy respondents for weighted Census Day persons was 5.5 percent for the original interview and 32.4 percent for the followup interview. The higher percentage of proxy for followup interviews is explained by the selection of followup cases largely upon proxy status in the original interview. Groups of persons that had percentages of proxy respondents significantly higher than other groups included the following. (The first percent shown refers to the Accuracy and Coverage Evaluation original person interview and the second percent shown refers to the person followup interview, unless otherwise noted.)

Age and Sex Groups:

- Young adult females aged 18-29 (8.0 and 37.5 percents),
- Young adult males aged 18-29 (7.2 and 37.0 percents),
- Adult males aged 30-49 (6.2 and 36.9 percents),
- Adult females aged 30-49 (5.7 and 34.3 percents),

Tenure and Residence Groups:

- Non-owners (10.5 and 45.1 percents),
- Persons in multi-unit dwellings (12.0 and 46.1 percents),
- Outmovers, in comparison to nonmovers (74.8 and 70.3 percents),
- Persons living alone (13.5 and 57.9 percents),

Geographic Groups:

- Persons in designated low census-return-rate areas (6.5 percent – original interview only),
- Persons in South (6.1 and 33.3 percents) and West (5.9 and 34.3 percents) regions,

Characteristics and Processing Status Groups:

- Persons with imputed characteristics (21.6 and 41.0 percents),
- Targeted Extended Search persons (8.4 percent – original Accuracy and Coverage Evaluation interview only), and
- Persons with imputed match status, i.e., unresolved search for a matching census enumeration (55.1 and 50.1 percents.)

Also, census-matched persons for whom answers to the original Accuracy and Coverage Evaluation interview questions were not the same as the census had higher percentages of proxy respondents. Those percentages ranged from 6.2 to 13.5 percent among items analyzed.

- Conversely, some household characteristic groups had statistically significant lower percentages of proxy response relative to other groups, including the following (Percents for the original and followup interviews are listed in parentheses as above.):
 - Parent or child of the person listed as reference person (3.8 and 24.4 percents) and
 - Persons in households of seven or more residents (2.5 and 16.4 percents.)
- In five variables investigated (Age/Sex, Tenure, Race and Hispanic Origin Domain, Type of Structure at Basic Street Address, and Household Size), persons with differing answers on Census 2000 relative to the A.C.E. had higher percents of proxy response. The difference in percentages of proxy response was particularly noticeable for Age/Sex where the proxy percent was only 3.6 when answers were consistent, but 13.5 when answers differed.
- Persons whose data were not supplied by a proxy in the original A.C.E. person interview had lower percents of item non-response (“Don’t know” and refusal responses as well as item missing) on the eight items of interest.
- Among proxy respondents, postal workers provided the highest percent of item non-response, except for Race and Hispanic Origin items, for which landlords provided the highest percents of item non-response.
- Outmovers from nonmover households had the highest percents of item non-response to Race and Hispanic Origin questions. Whole-household movers, both in-movers and out-movers, had higher percents of item non-response than non-movers for race (not Hispanic origin), but the opposite was true for Tenure. No general pattern of item non-response among mover groups was evident.

We recommend continuing efforts to minimize the early acceptance of proxy response as a guideline in data collection. We observed in many analyses that proxy response was related to unresolved or less complete answers to important questions. Although we cannot say that proxy response was the cause of less informed data, we suspect that it did contribute and that finding a more knowledgeable household respondent would make a positive difference.

We also recommend continued study of proxy data in future data collections to replicate or extend these findings.

We also recommend testing and evaluating alternatives to proxy data collection, such as use of administrative records. Administrative records would potentially reduce the need for proxy data and may yield more accurate household data in comparison to proxy data or imputed data.

REFERENCES

- Adams, T., Barrett, D., and Byrne, R. (2001). "Operational Plan for Accuracy and Coverage Evaluation (A.C.E.) for Census 2000," DSSD Census 2000 Procedures and Operations Memorandum Series S-TL-06, U.S. Census Bureau, Washington, D.C.
- Childers, D. (2001). "The Design of the Census 2000 Accuracy and Coverage Evaluation (A.C.E.)" DSSD Census 2000 Procedures and Operations Memorandum Series S-DT-01, Revised, U.S. Census Bureau, Washington, D.C.
- Fay, R. (1990). "VPLX: Variance Estimates for Complex Samples," Proceedings of the Section on Survey Research Methods, American Statistical Association, 266-271.
- Griffin, Deborah H. and Christopher L. Moriarity (1992) "Characteristics of Census Errors," *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 512-517.
- Haines, D. (2001). "Accuracy and Coverage Evaluation Survey: Computer Specifications for Person Dual System Estimation (U.S.) -Re-issue of Q-37," DSSD Census 2000 Procedures and Operations Memorandum Series Q-48, U.S. Census Bureau, Washington, D.C.
- Hill, Daniel H. (1987). "Response Errors in Labor Surveys: Comparisons of Self and Proxy Reports in the Survey of Income and Program Participation (SIPP)," *Proceedings of the Third Annual Research Conference*, Washington DC, Bureau of the Census, 299-319.
- Hogan, H. (1993). "The 1990 Post-Enumeration Survey: Operations and Results," Journal of the American Statistical Association, 88, 1047-1060.
- Killion, Ruth Ann (1999). "Outmover Tracing and Interviewing Evaluation," *Census 2000 Dress Rehearsal Evaluation Results Memorandum Series #C3*.
- Moore, Jeffrey C. (1988). "Self/Proxy Response Status and Survey Response Quality: A Review of the Literature," *Journal of Official Statistics Vol. 4, No.2, pp. 155-172*.
- O'Muircheartaigh, Colm (1986). "Correlates of Reinterview Inconsistency in the Current Population Survey (CPS)," *Proceedings of the Second Annual Research Conference*, Washington DC, Bureau of the Census, 208-234.

APPENDIX: Percent of Item Non-response for Proxy versus Non-Proxy

These tables represent an alternative to analyses in Tables 4.5.1 to 4.5.8. Proxy Type “Proxy” in the tables below represent a collapsing of “Landlord”, “Neighbor”, “Postal Worker”, and “Observation or Other” categories from the prior tables.

Table A.1. Percent of Item Non-response to the Tenure Item, by Proxy Status

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.5	1	2	<0.1	94.5
Proxy	2.8	2	1	0.2	5.5

Note: Criterion for levels to differ was $|t| > 1.645$

Table A.2. Percent of Item Non-response to the Age Item, by Proxy Status

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	2.1	1	2	<0.1	94.5
Proxy	7.4	2	1	0.3	5.5

Note: Criterion for levels to differ was $|t| > 1.645$

Table A.3. Percent of Item Non-response to the Gender Item, by Proxy Status

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.3	1	2	<0.1	94.5
Proxy	1.8	2	1	0.1	5.5

Note: Criterion for levels to differ was $|t| > 1.645$

Table A.4. Percent of Item Non-response to the Race Item, by Proxy Status

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	3.7	1	2	0.1	94.5
Proxy	13.6	2	1	0.7	5.5

Note: Criterion for levels to differ was $|t| > 1.645$

Table A.5. Percent of Item Non-response to the Hispanic Origin Item, by Proxy Status

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.7	1	2	<0.1	94.5
Proxy	9.9	2	1	0.6	5.5

Note: Criterion for levels to differ was $|t| > 1.645$

Table A.6. Percent of Item Non-response to the Kinship to Reference Person Item, by Proxy Status

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.2	1	2	<0.1	94.5
Proxy	1.7	2	1	0.1	5.5

Note: Criterion for levels to differ was $|t| > 1.645$

Table A.7. Percent of Item Non-response to the Other Residence Flag Item, by Proxy Status

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.6	1	2	<0.1	94.5
Proxy	12.5	2	1	0.4	5.5

Note: Criterion for levels to differ was $|t| > 1.645$

Table A.8. Percent of Item Non-response to the Special Place Flag Item, by Proxy Status

Proxy Type	Percent	Rank	Differs from	s.e.	n (percent)
Not a Proxy	0.5	1	2	<0.1	94.5
Proxy	8.7	2	1	0.4	5.5

Note: Criterion for levels to differ was $|t| > 1.645$