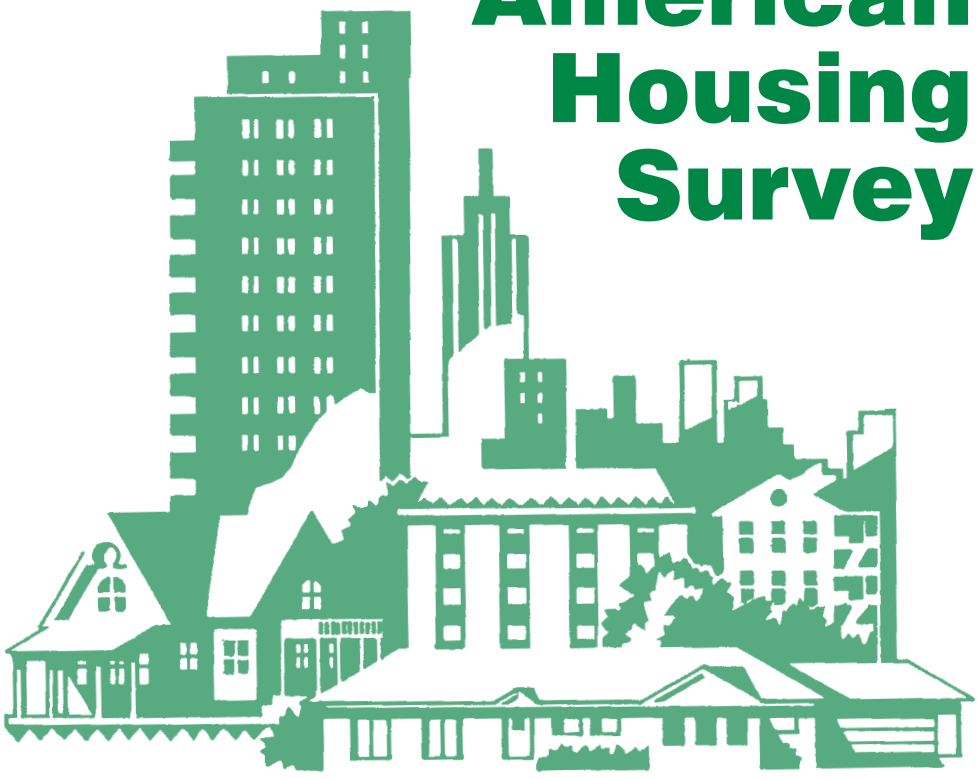
CURRENT HOUSING REPORTS

H121/95-1

American Housing Survey



A Quality Profile

by Rameswar P. Chakrabarty,
assisted by Georgina Torres

U.S. Department of Housing
and Urban Development
OFFICE OF POLICY DEVELOPMENT
AND RESEARCH

U.S. Department of Commerce
Economics and Statistics Administration
BUREAU OF THE CENSUS

Acknowledgments

This report was written in the Statistical Research Division (SRD), by **Rameswar P. Chakrabarty,** assisted by **Georgina Torres** of the Housing and Household Economic Statistics Division (HHES) under the general direction of **Lynn Weidman** of SRD. It was prepared as a result of the combined efforts of HHES, Demographic Statistical Methods Division (DSMD), Demographic Surveys Division (DSD), Field Division (FLD), and SRD. **James E. Hartman** and **Carol M. Mylet** of DSMD wrote the section, "Future Research/Planned Changes for AHS."

Edward D. Montfort, HHES; **James E. Hartman** and **Dennis J. Schwanz,** DSMD; **John C. Cannon,** DSD; and **Stephen T. Mann,** FLD provided invaluable support to the author by providing relevant source material, requested information, expertise, helpful comments, and suggestions, and reviewing the report for accuracy and completeness.

The author would also like to thank many other persons who provided relevant material on certain topics and/or reviewed some sections of the report. These persons are **John Bushery, Cornette Cole, Patricia Feindt, Andrea Meier, Irwin Schreiner,** and **Richard Summers,** DSMD; **Richard F. Blass, Medell Ford,** and **Richard Liquorie,** FLD; **Edward A. Hayes,** DSD; and **Barbara Williams,** HHES.

Maria Cantwell and **Carol Macauley** of SRD provided secretarial assistance in preparing the manuscript.

The staff of the Administrative and Customer Services Division (ACSD), **Walter C. Odom,** Chief, performed publication planning, design, composition, editorial review, and printing planning and procurement. **Barbara M. Abbott** provided publication coordination and editing.

CURRENT HOUSING REPORTS

H121/95-1

Issued July 1996

American Housing Survey

A Quality Profile

by Rameswar P. Chakrabarty,
assisted by Georgina Torres

U.S. Department of Housing and Urban Development

Henry Cisneros, Secretary

Michael A. Stegman,
Assistant Secretary for Policy Development and Research

U.S. Department of Commerce
Michael Kantor, Secretary

Economics and Statistics Administration
Everett M. Ehrlich, Under Secretary for Economic Affairs

BUREAU OF THE CENSUS
Martha Farnsworth Riche, Director

Economics and Statistics Administration

Everett M. Ehrlich,
Under Secretary
for Economic Affairs

BUREAU OF THE CENSUS

Martha Farnsworth Riche,
Director

Bryant Benton,
Deputy Director

Paula J. Schneider,
Principal Associate
Director for Programs

Nancy M. Gordon,
Associate Director for
Demographic Programs

Daniel H. Weinberg,
Chief, Housing and Household
Economic Statistics Division

OFFICE OF POLICY DEVELOPMENT AND RESEARCH

Michael A. Stegman,
Assistant Secretary for
Policy Development
and Research

Frederick J. Eggers,
Deputy Assistant Secretary
for Economic Affairs

Duane T. McGough,
Director, Housing and
Demographic Analysis Division

For sale by Superintendent of Documents,
U.S. Government Printing Office,
Washington, DC 20402

Contents

Tables

Tables—Con.

Tables—Con.

Figure

Exhibits

Chapter 1.
Introduction and Summary

INTRODUCTION

Objectives of the Report

This American Housing Survey (AHS) **Quality Profile** describes potential sources of errors in AHS data and quality control procedures used in the operation of the survey, and describes the magnitude of errors in AHS estimates. The description covers both sampling and nonsampling errors but emphasis is on the nonsampling errors. The report is intended to provide researchers and data users with a single source for a wide range of information on the quality of AHS data.

For information on sampling errors and their effect on analyses, refer to the appendix ''Source and Accuracy of the Estimates'' in published *Current Housing Reports* (H150 and H170 series).

The *Quality Profile* is intended for both data users involved in research or policy decisions and Census Bureau and U.S. Department of Housing and Urban Development (HUD) staff, (and others) who are responsible for or have an interest in AHS design and methodology. For the data users, this report describes the levels of error associated with specific categories of estimates that they use in their research or policy decisions. For those interested in the AHS design, data collection procedures, and estimation methods, this report describes the magnitude of errors associated with different features of the design, such as respondent rules, the interviewing process, and quality control procedures for data collection and processing operations.

This profile illustrates many design issues and methodological problems; some of them are unique to AHS but most are likely to arise in any household survey. Survey researchers concerned with survey design and methodological problems in surveys other than the AHS will also find many topics of interest in this report.

Sources of Data on Quality for AHS

AHS reports provide equations and tables that can be used to compute sampling errors. These sampling error functions and tables are produced by applying standard statistical techniques to the complex sample design of the survey.

The main sources of information on nonsampling errors are:

- Performance data such as coverage of the population, interview completion rates, item nonresponse rates, and results of reinterviews.
- Field and laboratory experiments designed to measure the effects on data quality of changing one or more features of survey design or procedures.
- Analytical studies involving statistical modeling which attempt to determine the size and direction of errors from individual sources of these errors.
- Comparison of aggregate data with similar data from other independent sources such as surveys, censuses, and administrative records.

Sources of Additional Information

Current Housing Reports (H150 series for the United States and H170 series for metropolitan areas) present tabulations and analyses of AHS data. Each report includes two appendixes that provide the following information:

Appendix A provides area classifications, definitions, and explanations of subject characteristics, and a facsimile of the AHS questionnaire.

Appendix B provides information on sample design, estimation, sampling errors and nonsampling errors. It also provides a set of standard error tables and illustrates the computation of standard errors for various types of estimates.

Appendix C. Beginning in 1993, the former appendix B has been divided into two separate appendixes (B and D). The new appendix B presents information on sample design and estimation. Appendix D describes the accuracy of the data.

The *Codebook for the American Housing Survey, Data Base: 1973-1993* (HUD and Bureau of the Census, 1990) provides information on sample design and errors in AHS (National and metropolitan) data.

Some papers on various aspects of AHS data quality have been presented and published in the proceedings of the annual meetings of the American Statistical Association. Most of the information in this report, however, comes from internal Census Bureau memoranda and documents. Readers interested in obtaining copies of any of these items should write to the Housing and Household Economic Statistics (HHES) Division, Bureau of the Census, Washington, DC 20233-8500 or call 301-763-8551.

Readers with questions about specific aspects of AHS design, methodology and data may contact:

Subject	Contact	Telephone
Survey design	Demographic Statistical Methods Division	301-457-1984
Estimation and weighting	Demographic Statistical Methods Division	301-457-1984
Sampling and nonsampling errors	Demographic Statistical Methods Division	301-457-1984
Data collection procedures	Field Division	301-457-1953
Data processing	Demographic Surveys Division	301-457-3873
Questionnaire design	Demographic Surveys Division	301-457-3877
Data characteristics and publications	Housing and Household Economic Statistics Division	301-763-8551

Structure of the Report

This quality profile describes each phase of the survey operations—sampling frame, survey design, sample selection, data collection, data processing, estimation, and data dissemination—and documents what is known about nonsampling errors associated with survey operations. When no data are available about the magnitude of a potential source of error and its impact on data quality, this is also indicated.

Chapter 2 provides a brief description of the AHS, its objectives, sample design, sampling frames, and sample size. More detailed information on various phases of survey operations and associated errors are described in subsequent chapters as follows:

Chapter 3. Data Collection Procedures
Chapter 4. Nonresponse Error
Chapter 5. Measurement Errors
Chapter 6. Data Processing
Chapter 7. Estimation
Chapter 8. Sampling Errors
Chapter 9. Comparison of AHS With Other Data

References are listed at the end of the report. .

SUMMARY

We describe design features of AHS (both the National and metropolitan area surveys), data collection procedures, data processing, estimation, variance estimation, and quality control procedures used in the operation of the survey . We discuss potential sources of errors associated with each survey operation, and document what is currently known about the magnitude of such errors. The discussion primarily focuses on nonsampling errors; information on sampling errors are provided in published Current Housing Reports. The quality profile unifies and summarizes available information on data quality from many reports and memoranda developed since survey inception in 1973. This part summarizes the main sources of nonsampling errors that affect the quality of AHS data, and the studies that attempt to measure their magnitudes.

Sample Design, Frames, and Undercoverage

The AHS is a stratified multistage probability sample of housing units. AHS-National's current sample design is based on the 1980 census. Each metropolitan area in the AHS-Metropolitan Sample (AHS-MS) has samples from the 1970 census and/or the 1980 census (see table 2.1). Most AHS-MS samples will be redrawn from the 1990 census.

The selection of housing units (HU's) within primary sampling units (PSU's) in the AHS-National or within metropolitan areas in the AHS-MS requires five separate non-overlapping sampling frames: (1) address enumeration districts (ED's), (2) area ED's, (3) special places, (4) new construction, and (5) coverage improvement. Frame development and sample selection within sample PSU's or metropolitan areas involve a complex system of automated and manual operations. For the area ED frame, a field operation—listing of addresses in sample blocks—is also necessary. All these operations are subject to errors. Some of the potential coverage problems are:

- Units constructed without permits in permit-issuing areas may be missed.
- If a permit is issued for a new structure at an existing address, that address may receive a duplicate chance of selection.
- Adequate coverage of mobile homes presents a variety of problems.

The magnitude of these coverage problems is not generally known, but is believed to be small in relation to the universe. Schwanz (1988a) estimated that undercoverage of mobile homes constructed after 1980 was close to 25 percent in the 1985 AHS-National.

Potential Sources of Errors in the Data Collection Procedure

The potential sources of nonsampling error in the AHS data collection procedure are many; for example, listing error, nonresponse, simple and correlated response variance, interview mode, difficulty in understanding questions, problems with year built, problems with multiunit structures,

etc. Some of these errors are systematically investigated and controlled as part of the AHS reinterview program. In this section, we discuss some sources of errors. Nonresponse and measurement errors are discussed in the sections ''Nonresponse Error,'' page 3, and ''Measurement Errors,'' page 4.

Listing error. Most of the listing errors occur in area segments. In 1988 a net error rate of -0.85 percent (standard error of 0.41 percent) in area segments in the Current Population Survey (CPS) reveals a slight undercount of units in the original listing (Waite, 1990c). An evaluation of listing errors for AHS is not available but their magnitude is likely to be similar to those of CPS. (See the section ''Listing by Observation in Area Segments,'' chapter 3, page 28.)

Problems with the coverage improvement screening procedure. The Coverage Improvement Screening Procedure does not always perform its intended function because the year in which the house was built is misreported by some respondents. The extent of coverage error due to misreporting of year built is not known. The response error in the year built data is discussed in the section ''Response error in year built data,'' see page 5.

Errors in Classification of Housing Units

The ''Building Loss—Vacant Other'' recheck program was conducted for AHS-National for several years to verify classification of all noninterview HU's (Type B and Type C) and all ''vacant-other'' units. (See the section ''Noninterviews,'' chapter 3, page 31, for definition of noninterview types.) Also, the coding of ''type of living quarters'' for HU's was checked. (See the section ''The '''Building Loss-Vacant Other'' Recheck Program for AHS-National,'' chapter 3, page 34 .)

The overall error rate for Type B was 7.1 percent in 1978 and 8.3 percent in 1979 (table 3.4). The overall error rate for Type C was 9.8 percent in 1978 and 11.6 percent in 1979 (table 3.5). The overall error rate in coding of ''type of living quarters'' by field representatives was 8.6 percent in 1977 and 5.4 percent in 1979 (table 3.6).

In 1979, the recheck program increased the AHS-National sample size by 137 units by including 275 units that were incorrectly deleted and dropping 138 units that were incorrectly retained. Thus it appears that at least some field representatives have a tendency to delete units that should not be deleted (classify true Type B's as Type C's) rather than include units that should not be included (classify true Type C's as Type B's), resulting in a decrease in the size of the sample.

During the first several years of the program, the information gained was used to modify the training, manuals, and data collection forms for losses to clarify some of the problem areas. As these improvements were incorporated it was felt that such a large-scale review that used large amounts of staff time was not needed. The last recheck, in 1985, was used to determine if the redesigned questionnaire had helped field representatives classify losses. The results of the 1985 recheck were encouraging and the program was not reinstituted for the 1987 survey.

Nonresponse Error

''Nonresponse'' refers to noninterview including unable-to-locate units and item nonresponse.

Unable-to-locate units. The units that cannot be found by field representatives are recorded as unable-to-locate (UTL) units. The UTL rates were less than 0.5 percent in address segments and exceeded 2 percent only in area segments in rural areas in the 1985 AHS-National. No data on UTL rates for AHS-MS are available but the rates are likely to be similar to the rates inside metropolitan areas (MSA's) for the National sample (table 4.1).

Noninterviews. The Type A noninterview rates were 4.2 percent, 3.2 percent, 4.2 percent, 4.4 percent, and 4.1 percent for the AHS-National in 1985, 1987, 1989, 1991, and 1993. (See the section ''Noninterviews,'' chapter 3, page 31, for definition of type A and table 3.2.)

In the AHS-MS, Type A noninterview rates varied by metropolitan area and by year. In 1986, Type A noninterview rate ranged from a low of 2.3 percent in Cincinnati to a high of 6.5 percent in San Antonio. Only 2 of 11 metropolitan areas had a noninterview rate above 5 percent. In 1987, Type A noninterview rate exceeded 5 percent in 4 of 11 metropolitan areas. In 1988, Type A noninterview rate was below 5 percent in all 11 metropolitan areas. In 1989, Type A noninterview rate ranged from a low of 2.4 percent in Minneapolis to a high of 9.1 percent in Washington, DC. The noninterview rate exceeded 5 percent in 7 out of 11 metropolitan areas in 1989. In 1990, Type A noninterview rate ranged from a low of 2.3 percent in Cincinnati to a high of 8.7 percent in Anaheim. The noninterview rate exceeded 5 percent in 3 out of 11 metropolitan areas in 1990. In 1991, Type A noninterview rate ranged from a low of 2.7 percent in St. Louis to a high of 8.5 percent in Northern New Jersey. The noninterview rate exceeded 5 percent in 2 out of 11 metropolitan areas in 1991. In 1992, Type A noninterview rate ranged from a low of 2.9 percent in Birmingham to a high of 5.6 percent in Norfolk. The noninterview rate exceeded 5 percent in 1 out of 8 metropolitan areas in 1992. In 1993, Type A noninterview rate ranged from a low of 4.6 percent in Detroit to a high of 8.4 percent in Washington, DC. The noninterview rate exceeded 5 percent in 3 out of 7 metropolitan areas in 1993. In 1994, Type A noninterview rate ranged from a low of 2.6 percent in San Diego to a high of 6.1 percent in Anaheim. The noninterview rate exceeded 5 percent in 3 out of 8 metropolitan areas in 1994 (see table 3.3).

Item nonresponse. Item nonresponse rates vary widely from item to item. Weidman (1988) estimated nonresponse rates for a set of 43 items for the 1985 AHS-National. Five items out of 43 had nonresponse rates greater than 10 percent and 12 had rates greater than 3 percent (table 4.2). If the questionnaire as a whole meets the minimum requirements for a completed interview, missing data for selected items are estimated by imputation (allocation). The imputed values are, at best, probabilistic in nature and subject to error, so potential biases from item nonresponse cannot be completely eliminated by imputation.

Measurement Errors

Nonsampling errors, other than coverage and nonresponse errors, that occur during data collection for a survey are called measurement errors. These errors may arise from different circumstances and causes (see the section ''Introduction, chapter 5, page 41). Information on some measurement errors are provided in this section.

Questionnaire design, content, and wording. It is well-known that questionnaire design; for example, wording of questions and order in which questions and possible response categories for a question are presented, affects responses. The 1985 questionnaire for AHS-National was finalized after receiving comments from regional offices and pretesting new questions in trial interviews to minimize response errors. As part of a continuing effort to improve the questionnaire, field representatives were requested to evaluate 1988 AHS-MS questionnaires and describe any problems on an evaluation sheet after they had completed the field work. Hayes (1989) recorded comments made by field representatives. These comments indicated that respondents had problems in understanding some questions. Need for better classification of buildings, basements, toilet breakdown, sewage breakdown, public/private water system, etc., were indicated. Studies of the ''reason-for-move'' over the years (see Montfort (1983a), Montfort (1983b), and Masumura (1981)) provide an interesting example of the development of a question over time and its impact on data quality (see the section ''Questionnaire Research and Development,'' chapter 3, page 26). As a result of the changes in the questionnaire in 1985 several items in the 1895-N and later are not comparable to similar data for 1973 through 1983. Items that changed on the 1985 questionnaire were: units in structure, rooms in unit, plumbing facilities, kitchen, and recent movers. A discussion of each item can be found under the topic of the same name in appendix C of the AHS- National report H150/93.

Currently, a thorough reevaluation of the questionnaire is underway as preparation for using computer-assisted interviewing for the AHS data collection beginning in selected metropolitan areas in 1996.

Interview mode. Interview mode—that is, personal visit, decentralized telephone interviewing, or computer-assisted interviewing—may affect the quality of data. Over the years the mode of interview has changed to some extent in the AHS-MS and considerably in the AHS-National.

For most of the history of data collection for AHS-MS, all cases, whether they were in sample for the first or a subsequent time, were interviewed in person. This has changed in recent years due to budget constraints. In the 1993 AHS-MS, interviews for cases that were in sample before and had telephone numbers were conducted over the telephone by decentralized telephone interviewing by field representatives with paper questionnaires.

In the AHS-National, telephone interviewing from an field representative's home became an acceptable alternative to personal interviewing as a result of the telephone experiments conducted in 1981 and 1983.

It is possible that Computer-Assisted Telephone Interviewing (CATI) techniques may collect data of even higher quality than achieved by face-to-face or telephone interviews. Using CATI may also help alleviate the effects of staffing retention problems in certain areas by reducing field workloads. Therefore, large-scale CATI experiments were implemented in conjunction with the 1987, 1989, and 1991 enumerations of the AHS-National sample to obtain information about the possible effects of CATI on the quality of AHS-National data. (See the section ''CATI experiments in the AHS-National, 1987, 1989, and 1991,'' chapter 5, page 41.) The results of these experiments are summarized below.

- The proportion of significant differences between CATI and non-CATI estimates were slightly higher than what would be expected due to chance alone. This indicated that the mode of interview affected the data (tables 5.1, 5.2, and 5.3).
- Nonresponse rates for CATI and non-CATI differed for certain items.
- There were differences in experience between CATI interviewers and field representatives.
- The 1991 moderate physical problems study revealed that CATI respondents underestimated deficiency items while non-CATI respondents overestimated them (tables 5.6 and 5.7).
- The gross difference rate analysis indicated that CATI had higher year-to-year change for some items and non-CATI for others. Neither CATI nor non-CATI estimates were generally better than the other for producing consistent responses (table 5.5).

Several changes were made in the 1991 CATI questionnaire and procedures to alleviate some of the factors that might have contributed to the differences. These changes had a positive impact on the 1991 results.

- There was a reduction in the overall proportion of differences between CATI and non-CATI estimates (table 5.2).

- There were substantial reductions in the CATI nonresponse rates for the items for which probes were added in the CATI questionnaire (table 5.4).
- The responses to certain items that rarely change were reconciled to improve the quality of data obtained in CATI interviews.

As a result of the 1991 CATI test results, it was decided to continue CATI use since it has many operational advantages. CATI can be used to monitor field representatives and reconcile questionable responses to improve data quality. In geographic areas with field representative retention problems, CATI can be used to reduce the field workload and to improve data quality.

Field representative effects. It is well-known that when a field representative collects data, his/her interaction with respondents and understanding or misunderstanding of questions can have important effects on the results. This is especially the case for questions that are subject to problems with definition or interpretation. All sample units surveyed by a field representative are subject to correlated field representative effects. This contributes ''correlated response variance'' to the total mean square error in the data.

There have been no formal interviewer variance studies in connection with AHS. However, the findings from other surveys and from censuses suggest that interviewer variance could be a significant source of errors for some items in the AHS (see the section ''Field Representative Effects,'' chapter 5, page 47).

Response errors. Response differences between interview and reinterview found in the AHS-National over the years are given in table 5.8 for selected items.

One percent of all households changed tenure. In particular, one percent of the owners were re-classified as renters, and 2 percent of the renters were re-classified as owners. The two interviews asked about tenure within 4 weeks of each other, so an actual change in tenure would be rare. The differences may be simple misunderstandings. They may also be ambiguous cases (such as property owned by a relative, which should be called rental). Note that response errors (as indicated by percentage of households changing answers between original interview and reinterview) increase with subjective items like street noise, traffic, etc.

Reinterview data can be used to obtain a statistical measure of discrepancies in responses called the 'index of inconsistency'. A summary of such indices computed from reinterview data from 1973 through 1985 has been compiled by Chakrabarty (1992a). Again, opinion questions like adequacy or inadequacy of recreation facilities, and items that are not easy to remember like the number of electrical blowouts in the last 90 days, have a high level of inconsistency (table 5.9).

Reinterview in the 1985 AHS-MS measured response variance of selected questions that generally fall into three categories: (1) major repairs, (2) mortgage, and (3) mobility. These three categories had moderate to high response variance as indicated by the index of inconsistency (tables 5.11, 5.12, 5.13, and 5.14).

Response error in year built data. Stating the year in which the structure was built has always been a problem for respondents in the AHS and other surveys; for example, CPS, and in the decennial census as well. This is particularly true when the respondent is not the first owner of the housing unit or is renting rather than buying.

A content reinterview for the 1980 census showed that the year built data have considerable response variance and bias (overreporting and underreporting). The multiunit structure data displayed higher response variability and bias than the single unit data. Also, the response variability in the year built data in the 1980 census was at about the same level as in the 1970 census (see, Bureau of the Census, 1986). Similar reinterview data from the AHS (National or MS) are not available.

The ''year built'' item was one of two items selected for a record check in the ''Tampa AHS Census Match Study'' (Tippett, 1988). The overall agreement of responses with the assessor's file was about the same for both census and AHS respondents. As expected, owners in the Census had better information on when the unit was built compared to renters. The high (14.5 percent) nonresponse rate for renters in this study for AHS might have biased the result. In any case, the differences between owners and renters based on a small AHS sample were not statistically significant (tables 5.15, 5.16, 5.17, and 5.18).

Young (1982) compared year built data for all housing units in the 1980 census and AHS. Several discrepancies existed between AHS and census estimates. A difference of 2.7 million units for the 1970-80 cohort was most striking (table 5.19). Young stated that, ''there are several possible reasons for the 1970-1980 cohort difference of 2.7 million units:

- A potential response error problem in the census. We know from past experience (1970 census evaluation program) that this is a problem.
- An excessive number of erroneous inclusions in the census; for example, duplicates, erroneous enumerations, etc. that were built during the period 1970-1980.
- Serious undercoverage problems in the AHS of units built during the period 1970-1980.''

Problems with the number of units in structure question. The number of units in a structure is a basic housing characteristic. A respondent is asked how many units there are in the structure in which his/her unit resides. A distinction is made between a housing unit; for example, an apartment, or townhouse, and the structure in which the

unit is contained. The structure or building may consist of one or many units. Furthermore, single unit structures are classified as either detached or attached to other structures. This question seems to give respondents a conceptual problem, especially in classifying townhouses, duplexes, and small attached units and in making a distinction between a housing unit and a structure.

Taueber, et al. (1983) compared 1980 census estimates of the totals of the ''units in structure'' categories with AHS estimates. The differences, except the totals, are greater than those expected from sampling error. Since the census was taken as of April 1, 1980 and the AHS date was around October 1980, due to interim new construction the total estimate of housing units was expected to be 800,000 to 1,000,000 units higher in the AHS than in the census. This is not the case however; the increase was only 335,000 units. The most notable difference existed in the ''5 or more units'' category (table 5.20).

Young (1982), who also examined the problem, stated that the possible reasons for this discrepancy were:

- ''Census misclassification error. There has been some concern that census respondents might have incorrectly identified certain types of single (or 2- to 4-unit structures) as 5-or-more-unit structures; for example, attached townhouses or garden apartments.''
- ''Serious undercoverage problems may exist in our current surveys for picking up new large multiunit structures.'' (See the section ''Problems With the Number of Units in Structure Question,'' chapter 5, page 58.)

Young (1982) also provided units in structure data separately for owners and renters (table 5.21).

The estimates for the number of 1-unit and 2- to 4-unit structures are remarkably close considering the time differential between the census and AHS. Most of the discrepancy in estimates is due to the ''5 or more units'' category. The AHS seems to have coverage problems for structures with 5 or more units and for within structure conversions.

The ''units in structure'' problem was also studied by Tippett (1988) in the Tampa AHS Census Match Study. The results (tables 5.22 and 5.23) further demonstrate the problem of classification in moderate- to large-sized buildings.

Finally, we considered a study described by Abernathy (1987) for the 1987 AHS-MS. The responses from Wave I of the Regional Office pre-edit were compared to the responses from the last enumeration period for AHS. (This is part of the continuing quality control program which checks for and corrects inconsistencies.) When the ''units in structure'' response is found to be inconsistent with the previous answer, the response is flagged.

The two main types of inconsistencies are as follows: ''units that were classified as one-attached one year and in a multiunit structure the other year; and units that were classified as in multiunit structures both years, but the number of units in the structure between survey years was inconsistent'' (table 5.24). Also, part of the quality control process was not only to detect the types of inconsistencies with the previous year, but also to check the corrected responses with the previous year. In other words, once the correction cycle is run on the data that are flagged as ''units in structure inconsistent,'' the responses are again checked with the entries from the previous enumeration period. At this point it has been determined that the majority of the corrected entries are consistent with the prior year's entries. Abernathy concludes, ''it appears that the pre-edit research is doing its job in reducing the classification problems that exist with the current year's data'' (see the section ''Problems With the Number of Units in Structure Question,'' chapter 5, page 58).

Problems with the tenure question. Tenure is important as a basic housing characteristic. The tenure question asks the respondent if he/she owns the unit, rents for cash, or occupies without payment of cash rent. The tenure question presents few conceptual problems for respondents, but the owner occupancy rates are persistently higher in surveys than in the census. This fact is documented by Taueber, Thompson, and Young (1983).

In the Tampa AHS Census Match Study (Tippett, 1988), the occupancy rate for owners in the AHS was 45 percent compared to 42 percent in the test census (table 5.25). Out of the 324 respondents who replied to both the test census and the AHS, 304 agreed and 20 gave conflicting responses (table 5.26). Thirteen of those twenty responses were reconciled. During the reconciliation reasons for the discrepancies were discovered and listed in the report as follows: ''for two cases, a change of tenure had occurred, so both were correctly enumerated; others resulted from mismarking of the item, different respondents, or a temporary interruption in the rent.'' These incidental discrepancies are not indicative of any problem that is inherent in the tenure question, and they do not help to explain the problem of the differences in the owner occupancy rates between the census and the AHS.

As an additional note, once the results have been reconciled the tenure item has an L-fold index of inconsistency in the low range, 11.08. This indicates that the respondents are answering the tenure question reasonably well.

Verification of reporting of cooperatives and condominiums. To evaluate the accuracy of the classification of housing units as cooperatives and condominiums in the AHS-National, part of the reinterview program for 1979 and 1983 focused on verifying responses to the AHS questions on cooperative and condominium status.

The verification followup showed that of the 1,634 units originally reported as condominiums, 62 (3.8 percent) were not condominium (table 5.27). And out of 196 units reported as cooperative in the original interview, 19 (9.7 percent) were verified to not be cooperatives (Hartnett, 1985).

These results reflect differences for only those housing units that were originally classified as cooperatives or condominiums. It is believed that the errors in the other direction are also a major source of the gross differences in reporting for these units. The regular reinterview program included questions on cooperative and condominium status for housing units not originally reported as a cooperative or condominium to provide an estimate of errors in the other direction. The results of this latter effort are not available.

Response error in multiunit structure characteristics. AHS field representatives have long reported that apartment dwellers often had little knowledge of the structural characteristics of their building. Fuels, heating equipment, and water supply were some of the affected items. For example, Smith (1985) analyzed the 1982 AHS-MS reinterview data and found both owners and renters showed moderate to high levels of inconsistency in reporting main heating equipment.

In order to evaluate the quality of responses from household respondents in multiunit structures and to investigate the feasibility of interviewing structure respondents, a multiunit structure (MUS) followup program was conducted with the 1984 AHS-MS.

In the MUS, Census Bureau interviewers asked a set of structure-related items (such as equipment and fuels) at all multiunit buildings in which there was a 1984 AHS-MS sampling unit. The MUS respondent was chosen to be knowledgeable about the entire structure in contrast to the AHS-MS household respondent who was to be knowledgeable about the specific unit.

A comparison of the AHS and MUS responses for the same building showed that the AHS apartment dwellers had a limited understanding of their building's characteristics. The amount of bias the AHS responses demonstrated were related to two aspects: first to the question being asked, and then, at a much lower level, to the size of the structure. Primary heating equipment was the most poorly reported item, while water source was the most consistently reported. The type of AHS respondent (that is, whether the respondent is the reference person, spouse, neighbor, or someone else) also affected the quality of the AHS data (see the section ''Multiunit Structures Followup to the 1984 AHS-MS,'' chapter 5, page 62).

The MUS followup was a one-time operation. As noted in Williams (1985) the MUS was relatively expensive for the amount of data improvement that resulted. Based on the results of the MUS followup, the AHS questionnaire items related to heating equipment were changed to improve the reporting for this item. There are no current plans to supplement or replace AHS household respondents' information with data from other sources.

Data Processing Errors

Data processing procedures for AHS-National and MS are essentially the same. Various phases of data preparation have built-in informal or formal quality control measures to minimize errors and to improve the quality of data. However, except for data keying, little quantitative information on errors in the different phases of data processing is readily available for AHS. Quality assurance results for keying and results of research on regional office pre-edit are summarized below.

Quality assurance results for keying. Statistical quality control methods are used to minimize data keying errors-(see the section ''Quality Assurance Results for Keying 1989 AHS-National,'' chapter 6, page 72). Results of keying verification are published regularly for AHS-National and AHS-MS. The national average incoming sample error rate was 0.16 percent for the 1989 AHS-National (table 6.1) and 0.19 percent for the 1989 AHS-MS (table 6.2). Incoming error rates and batch rejection rates for 100 percent verification for ''inexperienced keyers'' were higher than those from sample verification (tables 6.1 and 6.2). Note that the specified average outgoing quality limit (AOQL) for keying was 0.40 percent.

Research on regional office preedit. The regional office pre-edit is designed to improve the quality of the survey data. Data records (information as keyed from the control card and the questionnaires) are rejected if they fail to meet certain standards. Regional Office staff research the problems causing the records to be rejected, enter the corrective actions needed on the Correction Section of the Reject Listing, and key these corrections. For all metropolitan areas, the 1989 AHS-MS Regional Office pre-edit was conducted in four waves.

Abernathy (1991) analyzed Wave 1 reject data to (1) determine the status of the rejects, (2) determine the types of errors that caused the records to reject, and (3) compare the pre-edit reject corrections with how the reject situations would have been edited during the computer edit. The results are summarized below.

- There were 2,784 records that were rejected for 52 different reject reasons and about 90 percent of the total rejects were resolved.
- Seventy-seven percent of the total rejects were caused by specific data errors, 15 percent by relationship code errors, and 8 percent by other errors.
- The computer edit action was the same as the pre-edit action for fewer than half (45 percent) of the reject situations. However, for household demographic characteristics about 60 percent of the correction actions were the same as those the computer edits would have applied for these reject reasons (see the section ''Results of Research on Regional Offices Preedit for 1989 AHS-MS, chapter 6, page 73).

Comparison of AHS With Other Data

AHS data have been compared with census data to find differences in year built, units in structure and tenure items (see the sections ''Response Error in Year Built Data,''

chapter 5, page 56; ''Problems With the Number of Units in Structure Question,'' chapter 5, page 58; and ''Problems With the Tenure Question,'' chapter 5, page 60. In this section we provide comparisons of AHS utility costs with data from the Residential Energy Consumption Survey (RECS) and income data with independent estimates.

Comparison of AHS utility costs with RECS. RECS, conducted by the Department of Energy, collects utility costs data from utility company records. RECS data are, therefore, more accurate than AHS data provided by household respondents. A comparison of AHS utility costs with RECS data is provided in the codebook for AHS (HUD and Bureau of the Census, 1990). The results clearly show that AHS reports higher utility costs than the Residential Energy Consumption Survey. The discrepancy is fairly consistent over time, and also consistent for single-family detached homes. A plausible reason for the higher AHS figures is that households are more concerned about and, therefore, overemphasize high-cost months when they mentally average their bills for the AHS field representative.

The estimation of utility costs for AHS-National by regression using monthly utility cost data from the RECS public use file and some common RECS/AHS housing characteristics as independent variables was researched by Silwa (88a, 88b). Silwa (1989) provided specifications for deriving annual costs for electricity and natural gas. This method is now used to improve utility cost estimates for AHS. Another method used to improve respondent reporting is to include a request in the letter sent in advance to respondent households that they use records to determine utility costs for 4 specific months—January, April, August, and December.

Comparison of AHS income with independent estimates. It is well-known that income statistics derived from household surveys are generally biased due to response errors as respondents tend to underestimate income. A comparison of AHS income data with independent estimates of income (from national income and product accounts, the Social Security Administration, the Veterans Administration, etc.) and with the CPS is provided in table 9.1. The results show that the AHS estimates are lower than the independent estimates for total income and for every category other than self-employment income. The CPS estimate is also low but comes closer to the independent estimate. This may be largely due to the differences in income questionnaires and timing of CPS and AHS (March for the former versus the fall for the latter). Also, more detailed and extensive questions about income sources and amount by source are asked in CPS than in AHS. Finally, the CPS March supplement for income coincides with income tax time when respondents are more aware of nonwage incomes like interest, dividends, etc.

Recently, Williams (1992) provided an extensive comparison of the data on income that were collected in the 1989 AHS-National and the March 1990 CPS. This analysis at least partially supports the hypothesis that the AHS income estimates are lower than CPS largely due to the less detailed AHS income questions.

Future Research/Planned Changes for AHS

This section addresses several deficiencies mentioned previously. It contains actions we plan to take to correct some deficiencies as well as recommended research to help correct others.

Coverage. As noted in the section, ''Frames and Undercoverage,'' chapter 2, page 20, AHS is deficient in picking up mobile homes that are put in place after the census in address ED's. This is evidenced by the large undercoverage of new mobile homes compared to the Survey of Mobile Home Placements (SOMHP).

We currently plan to use the 1990-design National Health Interview Survey (NHIS) segment listings as a frame for picking up mobile homes that move to their current site. The 1990-design NHIS is a nationally-representative sample with an all-area design. This means NHIS will create listings of all housing units in address ED's. These listings will provide the frame for picking up these mobile homes that move to their current site. When doing the NHIS listings, information would be collected to help identify segments where there is a good chance of picking up these moved-to-site mobile homes in the future. AHS would update primarily these segments and possibly a subset of the other segments. There is also a possibility this frame could be used to pick up units in structures which converted from nonresidential to residential use in address ED's.

We also considered the SOMHP as a possible source to improve mobile home coverage. However, the SOMHP didn't have enough sample cases with good addresses and a frame based on the SOMHP would be more complicated and costly to implement. Since we're only interested in mobile homes in address ED's, the SOMHP mobile homes in area ED's would have to be identified and excluded. The SOMHP would also have to be modified to get better address information for us to use. Since the SOMHP currently doesn't need better address information, this could be costly. Also, in some areas, AHS may need a larger sample than the SOMHP can provide. (See the section ''Sample Design for AHS-National,'' chapter 2, page 12, for a discussion of the difference between address and area ED's).

Nonresponse error.

Household nonresponse. To determine how well interviewed housing units represent noninterviewed housing units, we plan to compare prior year or 1980 census characteristics of current year interviewed housing units and noninterviewed housing units.

Item nonresponse. AHS has items with high nonresponse that aren't currently adjusted for in the imputation procedure. A ''not reported'' category is included in the published AHS report for these nonresponses. We have made some questionnaire changes to reduce this problem. In addition, switching to a completely automated data collection system in 1997 should also help the response rate for some of these items. To further reduce the effect of item nonresponse, future research projects can focus on the following:

- Developing better ways to impute data for items we currently impute (for example, use regression analysis or check administrative records).
- Developing procedures to impute for items with a high level of nonresponse that we don't currently impute (for example, years on assumed mortgage, amount of mortgage assumed, amount mortgaged, monthly mortgage payment, purchase price of home).

Response error. AHS has many items with high response error, as noted in the section ''Response Errors,'' chapter 5, page 48, (for example, opinion of neighborhood and structure, water leakage) and items that erroneously change from year to year (for example, presence of a basement, mortgage). We plan to do several things to decrease the response variance associated with these items.

Administrative records. Certain items, such as year built and units in structure, are available from county or city tax offices. We are considering doing an administrative records check, like the Tampa records check of a sample of AHS cases to determine the magnitude of the problem for these items. (See the sections ''Response Error in Year Built Data,'' chapter 5, page 56, and ''Problems With the Number of Units in Structure Question,'' chapter 5, page 58). We will make a decision about what to do for the entire sample (for example, match all the cases to administrative records or compute an adjustment based on the results from matching a sample of records) based on the results from the administrative records check.

Dependent interviewing. Starting in 1997, a completely automated data collection system will be used for AHS. With this system, we will be able to use dependent interviewing for both personal and telephone interviews. Dependent interviewing uses responses from the prior interview as a check on the responses from the current year. We could not perform dependent interviewing accurately without moving to an automated data collection system.

One of the key elements for dependent interviewing is to first determine ''truth''(that is, the correct answer for a question). This will be done in 1997 by first asking the question and comparing it to the response from 1995. If they are the same, this answer will be considered ''true.'' If they are different, the respondent will be asked which answer is correct and this answer will be considered the ''true.''

Dependent interviewing will be used on three groups of items in the following ways:

- The first group are items which do not change from one year to the next, like year built. The first interview will determine ''truth'' for these items and they will never be asked again in future interviews.
- The second group are items which usually do not change but could change, like number of rooms. For these items the first interview will determine ''truth.'' In future interviews, the respondent will be asked if there was a change since the previous interview.
- The third group of items are items which could very well change, like tenure. For these items, the respondent will be asked the question at each interview. The answer will be reconciled if it differs from the previous one to determine which is correct.

Questionnaire changes. We have also changed the wording and placement of many of the items with high response variance like heating equipment, repairs and alterations, and water leakage to get more accurate responses.

Multiunit structures. In multiunit structures, respondents often do not know the correct answer for questions such as size of structure, year built, heating equipment, fuels, water, and sewage to name a few (see the section ''Multiunit Structures Followup to the 1984 AHS-MS, chapter 5, page 62). These questions could be asked of a more informed respondent such as the building manager, for example.

In 1995, we plan to collect some of the above multiunit structure information for rental properties containing AHS sample units in a special operation separate from the AHS. The answers from the informed respondent will be compared to the responses from the 1993 AHS-National sample units to determine the magnitude of the problem. We may also use this information as ''truth'' for dependent interviewing of the AHS-National sample cases. The results from this comparison will be used to decide if a structure-level respondent should be used for the AHS-MS.

Future response error measurement plans. After making changes to the questionnaire and switching to a completely automated data collection system, we plan to measure the effect these changes had on the response error and on the data.

We currently plan to measure response error for the 1996 AHS-MS. Part of the sample will be done by computer assisted personal interviewing (CAPI) with the changes to the questionnaire and part will be done without the questionnaire changes using a paper questionnaire. The response error from the two samples will be compared to determine the effect these changes had on response error and on AHS-MS estimates. In addition, estimates from the two samples will be compared to determine the effect these changes had.

In 1997, the non-CATI portion of the AHS-National sample will also switch to CAPI interviewing. Both the CATI and CAPI samples will use the new questionnaire. We plan to compare the estimates from 1997 to 1995 to see what effects the changes had on the data. However, this comparison will be somewhat tainted because of actual changes which could occur in that time period.

Chapter 2.
AHS Sample Design

OBJECTIVES OF AHS

The main objective of the American Housing Survey (AHS) is to provide a current, consistent, comprehensive, and accurate view of housing conditions and housing markets in the United States. The survey includes information on housing conditions, size and composition of the housing stock, and the characteristics of its occupants, information used as the basis for policy and program decisions by the U.S. Department of Housing and Urban Development (HUD). The survey also provides Congress, other Federal agencies, industry groups, and the research community with data used to assess housing adequacy and to make housing-related decisions.

HUD uses the AHS Metropolitan Survey to set maximum rent subsidy levels (Fair Market Rents) for its major housing assistance program, Section 8. Out-of-date data miss the effects of tight and loose market conditions, phenomena that are unique to individual housing markets, and their change over time. Failure to incorporate the latest rental market dynamics can result in inequitable or inefficient subsidy programs.

Policy analyses require knowledge of the current condition and affordability of the housing stock, and the current dynamics of housing markets upon which potential Federal policies operate. HUD annually compares its current level of housing assistance to the level of housing need, taking into account the condition, affordability, and usage of housing, based on the most recent AHS national data. The comparison serves as an indicator of housing program coverage and effectiveness.

Contemporary data requirements mandate a tight schedule for publication of AHS data. The AHS must provide a consistent view of the Nation's housing stock by maintaining standard definitions, data elements, and sample design through successive surveys during a decade. The AHS contains a broad and detailed data set addressing the significant policy issues associated with the nature and condition of the housing stock and the shelter experience of the Nation's households. The survey provides a sufficient range of data, on both a cross-sectional and longitudinal basis.

Accuracy goals are difficult to state in absolute terms. However, some key indicators of requirements are listed by HUD as follows:

1. To measure the 45th percentile of gross rents for 2 bedroom units of modest quality occupied in the last 2 years, with 95-percent confidence interval of plus or minus $20 per month, for the major metropolitan areas containing half of U.S. renters.
2. To identify changes in housing costs of at least 10 percent, for populations that are at least 5 percent of the Nation's housing population, with 95-percent confidence.
3. To identify changes in housing or neighborhood quality of the same magnitude for the same subgroups, with the same precision.
4. To identify changes in household composition affecting at least 1 percent of all households, with the same precision.

DESCRIPTION OF THE SURVEY

The AHS, conducted for the HUD by the Bureau of the Census, is actually two separate data collection efforts. One is a national sample and the other a metropolitan sample (MS). The AHS-National is a biennial survey of occupied and vacant housing units in the United States. The AHS-MS is a quadrennial survey of 44 large metropolitan areas, 11 per year, on average.

The surveys are conducted by field representatives who obtain the information from the occupants or, if the unit is vacant, from informed persons (landlords, rental agents, or knowledgeable neighbors). Interviews are conducted by personal visit or by telephone. The information reported by the field representative reflects the situation at the time of the survey, which is conducted during a 3 month period in the fall for the national and over 9 months for the MS. The Census Bureau conducted national surveys each year from 1973 to 1981. Beginning in 1983, the national survey is being conducted only in odd-numbered years. The MS surveys have been conducted every year, starting in 1974.

For the survey years 1973 through 1983, the data were collected for a sample of housing units located in the counties and cities that made up the 461 sample areas. A sample of housing units was selected in these areas from the 1970 census and updated by a sample of addresses from building permits to include housing units added after the census. Estimates of the counts and characteristics of the inventory were obtained for these sample units. The basic, designated sample consisted of approximately 60,000 housing units (HU's) located throughout the United States.

Beginning in 1985, a new, redesigned sample, selected from the 1980 census, has been used with a sample size of approximately 49,000 sample units.

On the questionnaires used for the AHS, the field representative records the information by marking a precoded check box or by writing in the entries. The information from the questionnaires is keyed directly to magnetic tape which is processed on the Census Bureau's computers through a number of editing and tabulating steps.

The AHS provides current information on the size and characteristics of the housing inventory and its occupants. Key statistics include tenure, the value and cost of housing, structural and equipment characteristics, housing quality indicators, household composition, race and ethnicity, income, and recent movers. These data are primarily used by HUD to establish Fair Market Rents and to measure housing inadequacy and the need for housing programs, by private industry to do market research, and by academic and private sector researchers to do other kinds of housing research.

During the first decade of the AHS, the overall design, purpose, and methodology did not change. This was particularly important, because the AHS is both a cross-sectional and a longitudinal survey. To derive the most benefits from the longitudinal data, year-to-year consistency was highly desirable. As different priorities emerged, HUD added new questions or subjects, such as neighborhood quality and energy consumption.

SAMPLE DESIGN FOR AHS-NATIONAL

The AHS-National is a stratified multistage probability sample of housing units. The current sample design is based on the 1980 census as outlined below.

Selection of Sample Areas

The United States was divided into areas made up of counties and independent cities referred to as primary sampling units (PSU's). Of these PSU's, 170 were known as self-representing (SR), since the sample from the PSU represented only that PSU. These 170 PSU's were in sample with certainty. The remaining PSU's were grouped into strata and were referred to as nonself-representing (NSR), since the sample of housing units (HU's) from the sample PSU represented all PSU's, both sample and nonsample, in the stratum. These NSR sample PSU's were selected in two steps.

1. The design for the Current Population Survey (CPS) involved strata consisting of one or more PSU's. Strata were formed independently within each State based on demographic and socioeconomic characteristics. In strata consisting of more than one PSU, CPS selected one of them to represent all PSU's in the stratum with probability proportional to the 1980 census population of persons 16 years of age or older.
2. To reduce costs, a subset of PSU's from CPS was selected. AHS field representative costs are greater in PSU's where CPS does not have any sample. Some of the CPS sample PSU's were self-representing (SR) for AHS. Those which were nonself-representing (NSR) were stratified independently within each region using characteristics from the CPS sample PSU's, weighted by the inverse of the probability of selection from CPS for that PSU, to represent the stratum from which it was selected. The characteristics which were used in stratifying the CPS sample PSU's were the following:

   1980 number of vacant housing units (HU's) for rent
   1980 number of owner-occupied HU's
   1980 number of occupied mobile homes or trailers
   1980 number of occupied HU's lacking some or all plumbing
   1980 number of occupied HU's with no complete kitchen facilities
   HU's built from 1970-1980
   1980 number of urban year-round HU's
   Population change from 1970-1980
   1980 number of owner-occupied HU's with value less than $25,000
   Heating degree days
   Cooling degree days

The 1980 number of HU's with a Black householder was also used in stratifying the South region. The 1980 number of HU's with a Hispanic householder was also used in stratifying the South and West regions. The last five characteristics listed above, as well as the Black and Hispanic householder where used, were given a weight twice as large as the other characteristics, indicating their greater importance. Of the CPS sample PSU's, 508 (SR and NSR) were grouped into 224 ''superstrata'' for AHS. For ''superstrata'' consisting of only one CPS sample PSU, that PSU was also selected for AHS. For ''superstrata'' consisting of more than one CPS sample PSU, one PSU was selected for AHS with probability proportional to the projected 1985 total housing unit count for the CPS stratum containing the CPS sample PSU.

Selection of the Sample Housing Units From the 1980 Census

The overall sampling rate used to select the sample of housing units from the 1980 census for the 1985 AHS was about 1 in 2,148. The within-PSU sampling rate was determined so that the overall probability of selection for each sample housing unit was the same (for example, if the probability of selecting a NSR PSU was 1 in 10, then the within-PSU sampling rate would be 1 in 214.8).

In census enumeration districts (ED's) where addresses were, for the most part, complete, and where new construction is monitored by permits (these ED's will be referred to as address ED's), all HU's from the 1980 census which received long-form questionnaires were stratified by the following characteristics:

CBUR (C = central city of an MSA, B = in urbanized area but not in the central city of an MSA, U = other urban, R = rural)
Tenure (owner, renter, vacant)
Number of rooms
Value (for owner-occupied units and vacant units for sale)
Rent (gross rent for renter-occupied units and contract rent for vacant units for rent)
Type of vacant (for all vacant except those for rent or for sale)

The stratification was done independently within CBUR within a region. A systematic sample of these units was selected at the rate of 2 in 2,148.

Group quarters (GQ's) are living quarters which do not meet the definition of a HU. A sample of GQ's was selected with probability proportional to the population in the GQ from the universe of GQ's which were classified as county homes; almshouses; poor farms; or soldiers', sailors', fraternal, or religious homes for the aged and were not known to have nursing care; and GQ's which were communes, rooming, boarding, or tourist homes. These GQ's were selected at the rate of 2 in 2,148. All other institutional GQ's were selected with equal probability and all other noninstitutional GQ's were selected with probability proportional to the population in the GQ. These other GQ's were selected at the rate of 2 in 3,069. This GQ sample was used to identify units in GQ's which had converted to HU's since the census.

For both the HU and GQ samples, one of every two units was assigned for interview for AHS, and the remaining units were assigned to the supplemental sample, some of which was to be used to increase the rural sample size every other survey year (that is, every fourth year) starting in 1987.

In ED's where at least 4 percent of the addresses were incomplete or inadequate, or where new construction was not monitored by building permits (most rural areas), a sample of 1980 census units which received long-form questionnaires was selected in several steps (these areas will be referred to as area ED's). ED's were stratified by the following characteristics:

CBUR
Median value of owner-occupied housing units
Number of children less than 6 years old
Total population age 65 or older
Number of owner-occupied HU's
Number of mobile homes or trailers
Number of units lacking some or all plumbing
Number of owner-occupied units with a value less than $45,000
Number of renter-occupied units with rent less than $200
Total minority (that is, Black or Hispanic) population
Number of one-room HU's

These stratifications were again done independently within CBUR. A sample of ED's was chosen with probability proportional to the 1980 census count of HU's and persons in GQ's combined in the following formula:

$$\left[\text{Number of HU's in the ED} + \frac{\text{Number of GQ persons in the ED}}{2.75}\right] / 4$$

A land area known as a segment was chosen within each sample ED. A sample of eight HU's which received 1980 census long forms was selected. If fewer than eight HU's received long forms, then all long-form recipients were selected and a sample of the remaining short-form questionnaire recipients were chosen so that a total of eight units were sampled. As was done for address ED's, the sample was selected so that the overall probability of selection for a unit was 2 in 2,148 and every other unit (four units per segment) was used in 1985. The remaining units were assigned to the supplemental sample which was to be used to increase the rural sample size every other survey year starting in 1987.

Selection of New Construction Housing Units in Permit-Issuing Areas

The sample of permit new construction was and continues to be selected from building permits issued such that the units were expected to be completed after April 1, 1980. For certain areas and structure sizes, this includes permits issued as early as March 1979, but, for the most part, includes only permits issued since July 1979. Only nonmobile home new construction is covered by the building permit frame. Within each PSU, building permits from each permit office were stratified by the Metropolitan Statistical Area (MSA) status of the office and chronologically ordered by month issued, so that the sample would be representative in terms of geography and month of issue. Compact (geographically proximate) clusters of approximately four housing units were created. These clusters were sampled at the rate of 8 in 2,148. Housing units in these clusters were subsampled at the rate of 1 in 4. One of every two sampled HU was assigned for interview for AHS with the remaining HU's assigned to the supplemental sample.

HUCS Sample

Housing units at addresses missed in the 1980 census or units which were at inadequately described addresses in the census address registers did not have a chance of

being selected for the AHS sample. A special study, done as part of the 1980 census, called the Housing Unit Coverage Study (HUCS), identified such units. A sample of the census misses in HUCS was included in the AHS sample. The probability of selecting these units was derived from the probability that they were included in HUCS.

Housing Units Added Since the 1980 Census

Non-newly constructed housing units added to the inventory since the 1980 census were represented using two methods. One method identified within-structure additions. These are units in structures which had a chance of being in sample because they contained at least one unit enumerated in the 1980 census. This method was used for the HUCS sample as well. The other method identified whole structure additions. These are units in structures for which none of the units was enumerated in the 1980 census.

In area ED's, all within-structure additions in structures containing at least one sample unit were interviewed for AHS. In address ED's, all within-structure additions in 1 to 15 unit structures containing at least one sample unit were interviewed for AHS. The probability of selection for these additions is the probability that any of the 1980 census units in the structure were selected for sample. In 16-or-more unit structures in address ED's, only units falling on AHS sample lines were interviewed for AHS. The probability of selection for these additions is the same probability of selection as for other sample units selected from the 1980 census (that is, 1 in 2,148).

In address ED's, whole structure additions were identified using area sampling methods. Under area sampling, all HU's within a land area are first listed and then a systematic sample is selected using a ''start with'' and ''take every'' so that a desired sample size is achieved based on the expected number of units within the segment. Segments from the National Health Interview Survey (NHIS), which were in sample in 1985, were used. Due to cost constraints, only NHIS areas which were in AHS PSU's or NHIS PSU's adjacent to AHS PSU's were used. Also, only units which were not already assigned to NHIS were eligible. A systematic sample of units not assigned to NHIS was selected within each land area. These units were then matched to the 1980 census address registers. If the address matched to the census, the unit was ineligible. (Only the basic address; that is, 801 Main Street, had to match; apartment number, mobile home site number, etc., did not have to match.) At the time of listing, eligible units were then screened further so that only units with no previous chance of coming into sample were picked up. (The screening eliminated units such as non-mobile home new construction, which is covered by building permits, and census misses.) The probability of selection for the units was the same probability of selection as for NHIS.

In area ED's where new construction is not monitored by building permits, all segments chosen for the sample in area ED's were used. An expected four units were chosen systematically within these segments to identify whole structure additions. Thus, the probability of picking up whole structure additions was 1 in 2,148. This sample was also screened at the time of listing using the same criteria as for address ED's. However, this sample was not matched to the census. One important difference to note is that new construction was not eliminated during the screening process.

In area ED's where new construction is monitored by building permits, only one-third of the segments chosen for the sample was used. An expected eight units were chosen systematically within these segments to identify whole structure additions. Thus, the probability of picking up whole structure additions was 1 in 3,222. This sample was also screened at the time of listing using the same criteria as for address ED's. Again, this sample was not matched to the census. Nonmobile home new construction was eliminated by the screening process since it is covered by the building permit frame.

SAMPLE SIZE—1985 AHS-NATIONAL

The basic sample was limited to approximately 49,000 units. A special group of 1,200 urban cases was chosen from the original 49,000 units to be the reference points (called kernel units) for an additional sample of neighboring units. This ''neighbor'' sample consisted of approximately the 10 housing units physically closest to the kernel units and was instituted to provide insights into the neighborhood dynamics that may affect the sample units. The neighbor sample cases are to be interviewed only in alternate survey years (1985, 1989, and so forth). In 1985, due to budget restrictions, only about 600 kernels out of 1200 were used, for selecting a sample of about 6,000 neighboring units. In 1989, about 900 kernels were used to select about 9,000 neighboring units. In the remaining survey years (1987, 1991...), as in the pre-redesign survey, a supplemental sample of rural housing units are to be interviewed along with the basic sample to derive better estimates for rural cases.

Interviewing of both neighbor sample cases and supplemental sample cases was discontinued in 1995 due to budget reductions.

SAMPLE DESIGN FOR AHS-MS

The American Housing Survey-Metropolitan Sample (AHS-MS) began in 1974 and continues to the present. Initially, the Census Bureau surveyed 20 standard metropolitan statistical areas (SMSA's) each year over a 3-year cycle, for a total of 60 areas. In 1977, in order to reduce costs, the survey was converted to a 4-year cycle—still with a total of 60 metropolitan areas. In 1984, the design was changed to include 44 metropolitan areas; 42 of the areas were part of the original 60 SMSA's and 2 areas (Tampa-St. Petersburg, FL MSA and Northern New Jersey area PMSA's) were new. These areas are listed in table 2.1 by year of interview. Interviewing normally takes place from April

through December. The Census Bureau has updated the geographic boundaries of the areas to agree with the 1983 Office of Management and Budget (OMB) definitions. Housing units from the 1980 census were chosen only for those counties added to comply with definitional changes between 1970 and 1983 and for the two new metropolitan areas. In 1987, a new sample from the 1980 census was drawn in the Houston, TX area PMSA's. Longitudinality with the previous sample is retained only for the original portion of each metropolitan area. The post 1994 sample design for the American Housing Survey Metropolitan Survey includes two new MSA's, Charlotte and Sacramento. Six metropolitan areas (New York, Northern New Jersey, Philadelphia, Detroit, Chicago, and Los Angeles) will no longer be included in the metropolitan survey. Supplemental sample will be added to these areas in the national survey and separate data published for these areas every 2 years.

The sample areas covered for metropolitan areas that remained in the AHS sample after survey year 1983 are consistent with the 1983 OMB definitions of a metropolitan statistical area (MSA), consolidated metropolitan statistical area (CMSA), or primary metropolitan statistical area (PMSA). In some instances, a given metropolitan area is a combination of primary metropolitan statistical areas and is referred to as a PMSA. In addition to adding new areas to some metropolitan samples in order to comply with the 1983 definitional changes, some new metropolitan sample areas have been added. Thus, each of the AHS-MS metropolitan areas will fall into one of three categories:

1. Areas of the same geographic area as defined for surveys prior to 1984 (areas in which the 1970 OMB definition of a SMSA is the same as the 1983 MSA, PMSA, or CMSA definition, 1970-based area).

2. Areas consisting of new area in addition to the 1970-based area.

3. Areas that are strictly 1980-based.

Table 2.1 shows the percent of the AHS-MS old construction sample that is 1970-based and 1980-based for each metropolitan area.

In 1984, the expected sample size in each metropolitan area was 4,250 housing units. In 1985, five large metropolitan areas—Detroit, Los Angeles-Long Beach, Philadelphia, San Francisco-Oakland, and Washington DC.—had expected sample sizes of 8,500 housing units and the other six smaller metropolitan areas had expected sample sizes of 4,250 housing units. In 1986, the expected sample sizes for larger metropolitan areas were reduced from 8,500 to 4,250 due to budget constraints. Thus, the expected sample size for each metropolitan area has been 4,250 housing units since 1986. Note that the sample in each metropolitan area was divided equally into nine random panels (panels 4 through 12 corresponding to planned month of interview). Certain panels were not interviewed in certain years to reduce costs (see table 2.1).

For the metropolitan areas interviewed in odd-numbered years, beginning in 1985 the number of sample cases for preparation of the publication tables is increased by combining national sample interviews with the regular MS interviews. This is possible because the national and MS questionnaires are basically the same; special weighting procedures are required.

Designation of AHS-MS Sample Housing Units

The sample housing units designated to be interviewed in a survey year consisted of the following categories, which are described below:

Housing units which were in the 1970-based area include:

1. All sample housing units that were interviewed in the previous survey.

2. All housing units that were selected as part of the 1976-1981 Coverage Improvement Program. These coverage improvement cases represented most of the housing units that, until these procedures were implemented, did not have a chance of selection.

3. All sample housing units that were type A noninterviews ((units eligible to be interviewed) in the previous survey.

4. All sample housing units selected from a list of new residential construction building permits issued since the previous survey. This sample represents the housing units built in permit-issuing areas since the previous survey.

5. All sample housing units that were added since the previous survey in sample segments from the nonpermit universe. This sample represents additions to the housing inventory since the previous survey in nonpermit-issuing areas.

6. In the 1970-based areas of the selected MSA's, all additional sample housing units selected from the 1980 Census of Population and Housing.

7. All sample housing units reinstated to sample. This includes units that had been dropped from sample due to sample reduction.

Housing units within new areas added to the metropolitan area in 1980 (1980-based area) include:

1. All housing units selected from the 1980 Census of Population and Housing.

2. All housing units that were selected from a list of new residential construction building permits. This sample represents the housing units built in permit-issuing areas since the 1980 census.

3. All sample housing units that were selected in sample segments from the nonpermit universe. This sample represents additions to the housing inventory in nonpermit-issuing areas since the 1980 census.

Table 2.1. **Metropolitan Areas in AHS-MS by Interview Years**

Metropolitan area	Old construction sample		Panels dropped, 1984 to 1995	
	Percent 1970-based	Percent 1980-based	1984	1988
Interview years: 1984 and 1988				
Birmingham, AL MSA	91.8	8.2	none	4
Buffalo, NY CMSA	100.0	0.0	none	4
Cleveland, OH PMSA	100.0	0.0	none	4
Indianapolis, IN MSA	100.0	0.0	none	4
Memphis, TN-AR-MS MSA	92.1	7.9	none	4
Milwaukee, WI PMSA	100.0	0.0	none	4
Norfolk-Virginia Beach-Newport News, VA MSA[1]	26.9	73.1	none	4
Oklahoma City, OK MSA	88.3	11.7	none	4
Providence-Pawtucket-Warwick, RI-MA PMSA's	93.2	6.8	none	4
Salt Lake City, UT MSA	83.4	16.6	none	4
San Jose, CA PMSA[2]	0.0	100.0	none	4
Interview years: 1985 and 1989	Percent 1970-based	Percent 1980-based	1985	1989
Boston, MA-NH CMSA	70.1	29.9	12	11 and 12
Dallas, TX PMSA	100.0	0.0	11 and 12	11 and 12
Detroit, MI PMSA	91.7	8.3	11 and 12	11 and 12
Fort Worth-Arlington, TX PMSA	96.2	3.8	11 and 12	11 and 12
Los Angeles-Long Beach, CA PMSA[1]	100.0	0.0	11 and 12	11 and 12
Minneapolis-St. Paul, MN-WI MSA	91.6	8.4	12	11 and 12
Philadelphia, PA-NJ PMSA	100.0	0.0	11 and 12	11 and 12
Phoenix, AZ MSA[1]	100.0	0.0	12	11 and 12
San Francisco-Oakland, CA area PMSA's[1]	100.0	0.0	11 and 12	11 and 12
Tampa-St. Petersburg, FL MSA[2]	0.0	100.0	12	11 and 12
Washington, DC-MD-VA MSA	93.3	6.7	11 and 12	11 and 12
Interview years: 1986 and 1990	Percent 1970-based	Percent 1980-based	1986	1990
Anaheim-Santa Ana, CA PMSA[1]	100.0	0.0	4 and 5	none
Cincinnati, OH-KY-IN PMSA	100.0	0.0	4 and 5	none
Denver, CO CMSA	97.6	2.4	4 and 5	none
Kansas City, MO-KS CMSA	91.0	9.0	4 and 5	none
Miami-Fort Lauderdale, FL CMSA[1]	63.3	36.7	4 and 5	none
New Orleans, LA MSA	95.2	4.8	4 and 5	none
Pittsburgh, PA CMSA	94.3	5.7	4 and 5	none
Portland, OR-WA CMSA	94.8	5.2	4 and 5	none
Riverside-San Bernardino-Ontario,CA PMSA[1]	100.0	0.0	4 and 5	none
Rochester, NY MSA	91.1	8.9	4 and 5	none
San Antonio, TX MSA	95.4	4.6	4 and 5	none
Interview years: 1987 and 1991	Percent 1970-based	Percent 1980-based	1987	1991
Atlanta, GA MSA	83.4	16.6	4 and 5	none
Baltimore, MD MSA	97.7	2.3	4 and 5	none
Chicago, IL area PMSA's	98.6	1.4	4 and 5	none
Columbus, OH MSA	80.4	19.6	4 and 5	none
Hartford, CT CMSA[1]	61.8	38.2	4 and 5	none
Houston, TX area PMSA's	0.0	100.0	4 and 5	none
New York-Nassau-Suffolk, NY area PMSA's	97.0	3.0	4 and 5	12
Northern NJ area PMSA's[2]	55.9	44.1	4 and 5	12
St. Louis, MO-IL CMSA	95.8	4.2	4 and 5	12
San Diego, CA MSA[1]	100.0	0.0	4 and 5	12
Seattle-Tacoma, WA CMSA	100.0	0.0	4 and 5	12

See footnotes at end of table.

Table 2.1. **Metropolitan Areas in AHS-MS by Interview Years—**Con.

Metropolitan area	Old construction sample		Panels dropped, 1984 to 1995
	Percent 1970-based	Percent 1980-based	1992
Interview year: 1992			
Birmingham, AL MSA	91.8	8.2	none
Cleveland, OH PMSA	100.0	o.o	none
Indianapolis, IN MSA	100.0	0.0	none
Memphis, TN-AR-MS MSA	92.1	7.9	none
Norfolk-Virginia Beach-Newport News, VA MSA[1]	26.9	73.1	none
Oklahoma City, OK MSA	88.3	11.7	none
Providence-Pawtucket-Warwick, RI-MA PMSA's	93.2	6.8	none
Salt Lake City, UT MSA	83.4	16.6	none
Interview year: 1993	Percent 1970-based	Percent 1980-based	1993
Boston, MA-NH CMSA	70.1	29.9	none
Detroit, MI PMSA	91.7	8.3	none
Minneapolis-St.Paul, MN-WI MSA	91.6	8.4	none
San Francisco-Oakland, CA area PMSA's[1]	100.0	0.0	none
San Jose, CA PMSA	0.0	100.0	none
Tampa-St. Petersburg, FL MSA	0.0	100.0	none
Washington, DC-MD-VA MSA	93.3	6.7	none
Interview year: 1994	Percent 1970-based	Percent 1980-based	1994
Anaheim-Santa Ana, CA PMSA[1]	100.0	0.0	12
Buffalo, NY CMSA	100.0	0.0	12
Dallas, TX PMSA	100.0	0.0	12
Fort Worth-Arlington, TX PMSA	96.2	3.8	12
Milwaukee,WI PMSA	100.0	0.0	12
Phoenix, AZ MSA[1]	100.0	0.0	12
Riverside-San Bernardino-Ontario,CA PMSA[1]	100.0	0.0	12
San Diego, CA MSA[1]	100.0	0.0	12
Interview year: 1995	Percent 1990-based		1995
Charlotte, NC MSA[2]	100.0		11
Columbus, OH MSA	100.0		11
Denver, CO CMSA	100.0		5, 7, 9, 11
Kansas City, MO-KS CMSA	100.0		11
Miami-Fort Lauderdale, FL CMSA[1]	100.0		5, 7, 9, 11
New Orleans, LA MSA	100.0		11
Pittsburgh, PA CMSA	100.0		11
Portland, OR-WA CMSA	100.0		11
San Antonio, TX MSA.	100.0		11

[1]100-percent permit-issuing in 1970 and 1980.
[2]New metropolitan area.

AHS-MS Original Sample Selection for the 1970-Based Area Sample of the Metropolitan Areas

The AHS-MS original sample for the 1970-based areas of the metropolitan areas, which, in 1970, were 100-percent permit issuing, was selected from two frames:

1. Housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the 1970-based permit-issuing universe).

2. Housing units constructed in permit-issuing areas since the 1970 census (the 1970-based new construction universe).

In addition, the sample for those metropolitan areas that were not 100-percent permit-issuing in 1970 included a sample selected from a third frame:

3. Housing units located in areas not under the jurisdiction of permit-issuing offices (the 1970-based nonpermit universe).

Sampling operations, described in the following paragraphs, were performed separately within the central city and balance of the metropolitan area, using the 1970 OMB definitions of the central city of each metropolitan area for each of the sample frames. The overall sampling rate used to select the sample for each metropolitan area was determined by the designated size of the sample. Each

metropolitan area had a sampling rate about the same for the central city and the balance, since the sample was distributed proportionately between the two, according to the corresponding distribution of total housing units.

Sample from the 1970-based permit-issuing universe. The major portion of the sample in each of the metropolitan areas was selected from a file that represented the 20-percent 1970 census long form sample of housing units enumerated in permit-issuing areas of the metropolitan areas during the 1970 Census of Population and Housing. This file contains records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied records were stratified by race of the head of household (non-Black/Black), and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to one of 50 strata according to its tenure (owner/renter), family size, and family income category as illustrated by the following table:

Family income	Tenure									
	Owner—family size					Renter—family size				
	1	2	3	4	5+	1	2	3	4	5+
Under $3,000										
$3,000 to $5,999										
$6,000 to $9,999										
$10,000 to $14,999 . .										
$15,000 and over										

Thus, the occupied housing unit records from the permit-issuing universe were assigned to one of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to one of the four vacant strata for either the central city or for the balance of a metropolitan area. A sample selection procedure was then instituted that would produce one-half of the desired sample. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the metropolitan area. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units of each of the special places were listed and subsampled at a rate that produced an expected four sample units, thereby insuring the necessary designated sample size.

Sample from the 1970-based new construction universe. The second frame from which the metropolitan area sample was selected was a list of new construction building permits issued since 1970 (the new construction universe). The sample selection from the list of new construction building permits was an independent operation within the metropolitan area. Using clerical procedures, the list of permits was stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate. In February 1984, the new construction sampling operation for the 1970-based and 1980-based areas were combined into one computerized system.

The universe sampled in the computerized system will be referred to in the estimation section as the 1980-based permit universe. Under these procedures, prior to sample selection, the list of permits was stratified by the date of issue, State, 1980 central city and balance, county or minor civil division, and permit office. Clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at twice the overall sampling rate. The housing units within each of the clusters were then subsampled so that two of the four housing units originally selected were kept in sample.

Sample from the 1970-based nonpermit universe. For those metropolitan areas that were not 100-percent permit-issuing, the remainder of the AHS-MS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection of a sample of census enumeration districts. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the metropolitan area. The probability of selection of an ED was proportional to the following:

$$\frac{\text{Number of housing units in 1970 census ED} + \dfrac{\text{Group quarters population in 1970 census ED}}{3}}{4}$$

The sample ED's were then divided into segments (small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units). At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units. The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census were included.

Sample Selection for the AHS-MS Coverage Improvement Program

The AHS-MS Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-Metropolitan area sample from the 1970-based permit-issuing universe and the 1970-based new construction universe within the 1970-based area. The coverage deficiencies included the following types of units:

1. New construction from building permits issued before January 1970, but completed after April 1, 1970.
2. Mobile homes placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For a detailed description of the coverage improvement sample selection process, see earlier reports in the H170 series for the years 1976 through 1981.

AHS-MS Sample Selection for the 1980-Based Area Sample of the Metropolitan Areas

The sample for new areas added to the 1970-based metropolitan areas, and metropolitan areas in sample for the first time that, in 1980, were 100-percent permit-issuing, was selected from two frames:

1. Housing units enumerated in the 1980 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the 1980-based permit-issuing universe).
2. Housing units constructed in permit-issuing areas since the 1980 census (1980-based new construction universe).

In addition, the sample for those metropolitan areas that were not 100-percent permit-issuing in 1980 included a sample from a third frame:

3. Housing units not under the jurisdiction of permit-issuing offices (1980-based nonpermit universe).

To satisfy confidentiality requirements in certain metropolitan areas, it was necessary to supplement the existing sample within the 1970-based area. The additional housing units were selected separately for each metropolitan area from the 1980-based permit issuing universe. Table 2.2 shows which metropolitan areas were 100 percent permit-issuing in 1970 and 1980.

Sample from the 1980-based permit-issuing universe. The major portion of the sample in each metropolitan area was selected from a file that represented all the housing units enumerated in permit-issuing areas during the 1980 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in group quarters. Sampling operations were done separately for noninstitutionalized group quarters and for all other housing units in permit-issuing areas. In addition, in order that an equal number of owner and renter housing units were selected in each metropolitan area, a selection rate that differed by tenure group was used. Before the sample was selected, the housing units that were not classified as group quarters were stratified into 60 categories by tenure, contract rent, value, and number of rooms as illustrated by the following table:

Contract rent and value	Number of rooms		
	1 to 3	4 to 5	6 or more
RENTER			
Contract rent:			
Less than $100			
$100 to $149			
$150 to $199			
$200 to $249			
$250 to $299			
$300 to $349			
$350 to $399			
$400 or more			
Not available			
OWNER			
Value:			
Less than $20,000			
$20,000 to $29,999			
$30,000 to $34,999			
$35,000 to $39,999			
$40,000 to $49,999			
$50,000 to $64,999			
$65,000 to $79,999			
$80,000 to $99,999			
$100,000 to $149,999			
$150,000 or more			
Not available			

The group quarters housing units were grouped into two strata: institutionalized group quarters; and noninstitutionalized group quarters.

The following sample selection procedures were then implemented separately within the central city and balance of the metropolitan area. All units were sorted by the 1980 central city and balance, stratum, State, district office, ED, and census serial number. The sample selection procedure was then implemented separately for (a) institutionalized group quarters and non-group quarters housing units, and (b) non-institutionalized group quarters.

Individual housing units were selected for the non-group quarters but each institutionalized group quarters had one chance of selection. Before the sample selection for the noninstitutionalized group quarters was implemented, the following measure of size was calculated for each group quarter record:

$$\frac{(1/4) \times (\text{Total group quarters population})}{2.75}$$

The noninstitutionalized group quarters were then selected proportionate to this measure of size.

Sample from the 1980-based new construction universe. The second frame from which the metropolitan area sample was selected was a list of new construction building permits issued since 1980 (the new construction universe). The sample selection from the list of new construction building permits was an independent operation within each metropolitan area. This operation was described in the discussion of the 1970-based new construction universe.

Sample from the 1980-based nonpermit universe. For those metropolitan areas that were not 100-percent permit-issuing, the remainder of the AHS-MS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (the 1980-based nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection of a sample of census ED's within these areas (using the overall sampling rate). Prior to this sample selection, the ED's were sorted by State, district office, and enumeration district number. The probability of selection of an ED was proportionate to the following:

$$\frac{\text{Number of housing units in 1980 census ED} + \dfrac{\text{Noninstitutionalized group quarters population in 1980 census ED}}{2.75}}{4}$$

The sample ED's were then divided into segments (small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units). At the time of the survey, those segments that did not have an expected size of four housing units were further subdivided to produce an expected four sample housing units. Following the division, a segment from each sample ED was selected. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1980 census as well as housing units built since the 1980 census are included.

FRAMES AND UNDERCOVERAGE

The selection of HU's within PSU's in the AHS-National or within metropolitan areas in the AHS-MS requires five separate non-overlapping sampling frames: (1) address ED's, (2) area ED's, (3) special places, (4) new construction, and (5) coverage improvement. Frame development and sample selection within sample PSU's or metropolitan areas involve a complex system of automated and manual operations. For the area ED frame, a field operation—listing of addresses in sample blocks—is also necessary. All these operations are subject to errors.

A comprehensive account of quality control procedures and information about errors associated with frame development and sample selection based on the 1970 census is given by Brooks and Bailar (1978). Although their focus was on CPS, their findings apply equally to the AHS, which was based on the same set of frames and sampling procedures. They identify several potential coverage problems associated with the address frames used.

- Units constructed without permits in permit-issuing areas may be missed.
- If a permit is issued for a new structure at an existing address, that address may receive a duplicate chance of selection.
- Adequate coverage of mobile homes presents a variety of problems.

The magnitude of these coverage problems is not generally known, but is believed to be small in relation to the universe.

The redesigned 1985 AHS-National sample is based on frames developed from the 1980 census. The current Sample in the AHS-MS is based on both 1970 and 1980 censuses. (See the section ''Sample Design for AHS-MS,'' page 14.) The frame development procedures used in the 1980 census were quite similar to those used to develop frames from the 1970 census. One difference, aimed at coverage improvement, was that the percentage of complete addresses required for an ED to be included in the address ED frame was increased from 90 to 96 percent for the 1980 census.

Another area of the AHS sample where coverage deficiencies exist is the sampling of building permits to represent conventional (nonmobile home) new construction. Due to time constraints, only permits issued more than 6 months before interviewing began are eligible to be selected to represent conventional new construction. This is more of a problem for single-unit rather than multiunit structures. In fact, the time lag between issuance of a permit and completion of construction for multiunit structures is generally more than six months depending on the size of the structure. Also, new construction in special places such as colleges or military bases is not covered. This is a deficiency in both permit and nonpermit areas.

Schwanz (1988a) estimated that undercoverage of mobile homes constructed after 1980 was close to 25 percent in the 1985 AHS-National. Coverage of new mobile home parks in address ED's was very poor.

Some coverage problems arise when permit issuing offices change the boundaries of the area within their jurisdiction or discontinue issuance of permits. If an area in the area ED frame is brought under the jurisdiction of an existing permit office, new units in that area will have a duplicate chance of selection, through the area ED frame and the new construction frame. Conversely, if an existing permit office stops issuing permits, new units in the areas which it covers will have no chance of selection. As of mid-1988, it was estimated that about 120 new housing units per month in such areas (equivalent to 0.08 percent of the total newly-constructed units authorized for the entire country) had no chance of selection (Loudermilk, 1989).

In identifying whole structure additions in address and area ED's, units which were in sample were screened to see if they were eligible for interview. The screening operation involved asking a series of questions. Therefore, the quality of coverage in these areas is only as good as the quality of the responses to these questions. It is conceivable that eligible units were omitted and ineligible units were included because the respondents' answers to the screening questions were incorrect. In addition, the quality of the listing of addresses will also affect the coverage of whole structure additions.

It is also believed that a coverage deficiency exists for units which were nonresidential at the time of the 1980 census, but have since converted to residential units. The magnitude of this deficiency is not known.

The ratio estimation procedures are used to correct HU coverage deficiencies in both AHS-National and AHS-MS.

The proportion of sample addresses in the 1985 AHS-National that came from each of the five frames are given in table 2.2.

It can be seen that 65.2 percent of sample addresses came from address ED's, 25.6 percent from area ED's, and 6.9 percent from the new construction frame in the 1985 AHS-National. Table 2.3 presents the distribution of sample addresses by frame for AHS-MS in 1988 and 1989.Metropolitan areas naturally had a higher proportion of the sample from new construction but a lower proportion from area ED's compared to the AHS-National.

Table 2.2. **Distribution of 1985 AHS-National Sample Addresses by Frame**

Frame	Percent of addresses
Address ED's	65.2
Area ED's	25.6
Special places	0.4
New construction	6.9
Coverage improvement	1.9

Table 2.3. **Distribution of Sample Addresses by Frame for AHS-MS 1988 to 1992**

Metropolitan area	Percent of addresses—frame			
	Address ED's[1]	Area ED's	New construction	Coverage improvement
1988				
Birmingham, AL MSA	57.6	18.0	23.0	1.4
Buffalo, NY CMSA	84.9	1.3	13.0	0.7
Cleveland, OH PMSA	82.4	0.5	16.0	1.1
Indianapolis, IN MSA	65.3	5.3	27.9	1.5
Memphis, TN-AR-MS MSA	61.0	3.8	33.6	1.5
Milwaukee, WI PMSA	76.9	0.2	22.1	0.9
Norfolk-Virginia Beach-Newport News, VA MSA	71.9	0.0	27.3	0.8
Oklahoma City, OK MSA	51.8	7.9	39.1	1.1
Providence-Pawtucket-Warwick, RI-MA PMSA's	76.6	0.0	21.6	1.8
Salt Like City, UT MSA	55.2	0.0	42.1	2.7
San Jose, CA PMSA	89.5	0.0	10.5	0.0
1989				
Boston, MA-NH CMSA	77.2	0.1	15.1	7.6
Dallas, TX PMSA	44.3	8.3	45.3	2.2
Detroit, MI PMSA	79.2	0.0	17.1	3.7
Fort Worth-Arlington, TX PMSA	43.1	10.1	43.5	3.2
Los Angeles-Long Beach, CA PMSA	75.0	0.0	20.6	4.3
Minneapolis-St. Paul, MN-WI PMSA	65.2	0.5	31.2	3.0
Philadelphia, PA-NJ PMSA	75.0	0.7	15.9	8.4
Phoenix, AZ MSA	31.1	0.0	62.8	6.1
San Francisco-Oakland, CA area PMSA's	70.1	0.0	21.4	8.5
Tampa-St. Petersburg, FL MSA	75.4	0.0	24.6	0.0
Washington, DC-MD-VA PMSA	65.3	0.1	29.8	4.8

See footnote at end of table.

Table 2.3. **Distribution of Sample Addresses by Frame for AHS-MS 1988 to 1992**—Con.

Metropolitan area	Percent of addresses—frame			
	Address ED's[1]	Area ED's	New construction	Coverage improvement
1990				
Anaheim-Santa Ana, CA PMSA	50.0	0.0	47.2	2.8
Cincinnati, OH-KY-IN PMSA	72.9	0.5	23.7	2.9
Denver, CO CMSA	51.3	0.0	46.4	2.3
Kansas City, MO-KS CMSA	64.7	6.6	26.5	2.2
Miami-Ft. Lauderdale, FL CMSA	61.3	0.0	36.0	2.7
New Orleans, LA MSA	62.8	9.0	26.1	2.1
Pittsburgh, PA CMSA	79.6	7.9	10.8	1.9
Portland, OR-WA CMSA	60.4	0.6	35.8	3.2
Riverside-San Bernardino-Ontario, CA PMSA	44.3	0.0	49.1	6.7
Rochester, NY MSA	72.0	7.0	19.4	1.6
San Antonio, TX MSA	51.0	17.1	30.8	1.1
1991				
Atlanta, GA MSA	47.1	0.9	49.7	2.2
Baltimore, MD MSA	63.5	0.3	34.3	1.9
Chicago, IL area PMSA's	62.2	4.9	26.4	6.5
Columbus, OH MSA	61.7	4.3	32.3	1.7
Hartford, CT CMSA	75.7	0.0	22.7	1.5
Houston, TX area PMSA's	68.0	3.7	28.3	0.0
New York-Nassau-Suffolk, NY PMSA's	63.3	0.4	33.6	2.7
Northern NJ PMSA's	75.2	0.9	21.6	2.3
St. Louis, MO-IL CMSA	67.3	5.6	25.1	2.0
San Diego, CA MSA	43.7	0.0	53.6	2.7
Seattle-Tacoma, WA CMSA	56.3	0.2	40.5	3.0
1992				
Birmingham, AL MSA	62.3	20.8	16.4	0.6
Cleveland, OH PMSA	87.9	0.5	11.6	0.0
Indianapolis, IN MSA	76.3	5.2	18.0	0.4
Memphis, TN-AR-MS MSA	72.9	4.8	21.9	0.4
Norfolk-Virginia Beach-Newport News, VA MSA	94.6	0.0	5.2	0.2
Oklahoma City, OK MSA	59.8	9.3	29.7	1.1
Providence-Pawtucket-Warwick, RI-MA PMSA's	86.1	0.0	13.5	0.4
Salt Lake City, UT MSA	71.3	0.0	27.6	1.1

[1]Includes special places.

Chapter 3.
Data Collection Procedures

INTRODUCTION

This chapter describes data collection activities, costs, and quality control procedures, and their impact on data quality. Detailed information about all features of data collection activities are given in the Field Representatives' Manual (Bureau of the Census, 1985a), which is periodically updated.

DATA COLLECTION STAFF

The data collection staff works from 12 regional offices (RO's) under the supervision of a regional director and under the overall supervision of the Chief of the Field Division. The AHS is the responsibility of the demographic program coordinator who has an AHS program supervisor on his/her staff. Each RO has a crew of field representatives and a staff of supervisory field representatives who assist the AHS program supervisors in on-the-job training, observation, and reinterview programs. Each field representative is a part-time employee who works out of his/her home. The formal titles of Census Bureau field interviewers and field supervisors have recently been changed to field representatives and senior field representatives, respectively.

Interviews of sample housing units take place during the summer and fall of the survey year. For example, data collection began in July 1987 and continued through December 1987 for the 1987 AHS-National. Information for the 1990 AHS-MS was collected by interviewers from June 1990 through November 1990. The information reported for a given unit reflects its situation at the time that unit is interviewed.

MODE OF INTERVIEW

At present, the telephone is the preferred mode of interview and the one used in an increasing number of AHS-National interviews. In the 1987 AHS-National, one-third of the sample was assigned to computer-assisted telephone interviewing (CATI) at a centralized location and the remaining two-thirds were assigned for the field. Field representatives were allowed to conduct interviews by telephone from their home for households that were in sample before; they had telephone numbers listed on control cards. About 56 percent of the occupied units were interviewed by personal visits in 1987 and 52 percent in 1989. All the AHS-MS cases, whether they were in sample for the first time or a subsequent time, were interviewed in person. The rules have changed in recent years due to budget constraints. The new rule on AHS-MS for first-time households is that if the field representative goes to a household three times and cannot find a respondent, but somehow gets the name and phone number, the field representative can conduct a telephone interview. In the past, if a field representative needed to conduct a telephone interview he/she had to obtain permission from the program supervisor. In the 1993 AHS-MS, interviews for cases that were in sample before and had telephone numbers, can be conducted over the telephone. The impact of interview mode on data quality is discussed in chapter 5.

FIELD REPRESENTATIVE CHARACTERISTICS, TRAINING, AND SUPERVISION

Field Representative Characteristics

Field representatives and CATI interviewers are generally part-time employees. Field representatives usually visit respondents' homes. CATI interviewers telephone from offices in Hagerstown, MD and Tucson, AZ. The AHS-National uses mostly experienced field representatives but AHS-MS typically hires new field representatives as each metropolitan area is covered only once every 4 years. Field representatives are considered experienced if they have worked on the AHS in a prior enumeration or are currently working on another Census Bureau demographic survey. Experienced supervisory statisticians assign and recruit staff for the AHS. Both surveys use crew leaders to assist the supervisor in observation, reinterview, and Type A followup. Senior crew leaders may assist in recruiting and training while both senior and junior crew leaders may take up emergency assignments for interviewing on short notice. The average workload is from 30 to 50 households per month for both AHS-National and AHS-MS. Field representatives were paid between $7.00 and $8.70 per hour in 1993.

Field Representative Training

The training for AHS field representatives includes a home study, classroom training, special topics self-studies, and on-the-job training.

Initial training. New field representatives receive intensive training, including one day of advance self-study followed by 3-1/2 days of classroom training. Training sessions include: lectures, audio-visual presentations, several mock interview exercises, and discussions. Trainees receive detailed information on their jobs, the concepts and definitions used in the survey, and specific interviewing techniques, such as probing. As part of the initial training, a supervisor or supervisory field representative observes each new field representative during his/her first 2 or 3 days of interviewing. Experienced field representatives receive 1 day of advance self-study followed by 1 day of classroom training.

Supplemental training. Field representatives found to be weak in certain aspects of the survey, such as completion rates and accuracy, are given supplemental training to help them meet the Census Bureau's standards.

Quality Control Procedures and Supervision

To ensure completeness and accuracy, field representatives conduct an edit on their AHS questionnaires. They look for inconsistent and missing entries. As appropriate, they call back the household to obtain the correct information. To keep costs down, these followup contacts are completed over the telephone. The field representative's work is monitored and feedback provided to them in several ways:

Questionnaire checks. Completed questionnaires are sent to the Census Bureau's RO's, where they are subjected to simple computer edits that are incorporated into the data-entry programs. A sample of questionnaires receives a complete clerical edit. More complex edits are performed on the computerized records when they are received at Census headquarters. In some instances, field representatives may be contacted to resolve problems identified in these edits.

Performance observation and standards. Field representative performance is measured by observation and reinterview results, accuracy rates, response rates, and production rates. The program supervisor has the responsibility for reviewing all observation reports to ensure that the observation was completed as required. The observation program provides on-the-job training, motivates the field representatives to become more efficient and effective employees, and provides supervisory personnel with a better insight into the field representatives' working conditions. If a field representative's performance problems prompted an observation, the observation must be directed toward solving them. If the problems were solved or if additional attention is still required, the program supervisor is made aware of this through discussions with the observer and through the written observation reports. An analysis of a field representative's performance report will indicate whether the field representative has made an excessive number of errors while completing the questionnaires. This is determined through the clerical edit of a sample of completed questionnaires. Each time an edit is performed an accuracy rate is calculated.

Accuracy rate standards are as follows:

	Percent
Outstanding	98.51 or more
Commendable	96.01 to 98.50
Fully successful	86.01 to 96.00
Marginal	81.51 to 86.00
Unsatisfactory	81.50 or less

Response rates (percent of sample households for which some response is obtained) are calculated.

The response rate standards are as follows:

	Percent
Outstanding	99.1 or more
Commendable	97.1 to 99.0
Fully successful	95.1 to 97.0
Marginal	92.1 to 95.0
Unsatisfactory	92.0 or less

Production rates, based on the time spent on each case, are calculated for each field representative.

The hours per case (production) standards are based on these figures:

	Percent
Outstanding	1.0 or less
Commendable	1.01 to 2.00
Fully successful	2.01 to 3.25
Marginal	3.26 to 4.5
Unsatisfactory	4.51 or more

Attitude toward the job, as shown by prompt attendance at classroom training sessions, prompt and satisfactory completion of self-study exercises, and the adherence to assignment deadlines, also is included in the overall evaluation of the field representatives. The accuracy, response, and hours per case standards are also tools. These are used in conjunction with observation, reinterview results, and a general knowledge of each field representative's attitude. The production (hours per case) is intended as a guide only and poor performance in this area is not used *exclusively* to remove an employee. If an field representative's performance is not acceptable, the program supervisor takes steps to help the field representative to improve his/her work.

Accuracy and consistency. The regional office preedit operation is a systematic process to check the completeness and accuracy of the keyed data in the regional office. The regional office can correct errors found during this operation before the data enters the final computer edit and allocation operations. The regional office preedit passes

the keyed data through a program that checks selected key data items for completeness and accuracy. Records that have one or more ‘‘problems’’ are listed (rejected) so the office can review the data and make any corrections necessary. This operation was designed to replace the much more labor intensive, time consuming, and error prone clerical edit of these items. The program also checks for consistency with prior year data concerning type of living quarters and number of units in structure so that errors in these two key items are reduced and the quality of the longitudinal data is enhanced.

Reinterviews. A systematic reinterview program serves the dual purposes of checking a sample of the work of individual field representatives and identifying aspects of the field procedures which may need improvement. Reinterviews are completed as soon as possible after the original interview, and are usually conducted on the telephone by supervisory field representatives or other members of the supervisory staff. The reinterviews are used to determine whether the field representatives visited the correct units, classified noninterviews correctly, and determined household composition correctly. In addition, several questionnaire items are checked to verify that the field representative asked these items during the original interview. The results of the reinterviews are used to take corrective action, such as supplemental training and observation for field representatives whose work is below standard.

DATA COLLECTION INSTRUMENTS

The primary data collection instruments for AHS-National and AHS-MS are the control card and the questionnaire. For each sample address, a control card is completed at its first enumeration year (for example, 1985 for the redesigned national sample) and updated at each subsequent enumeration. At the first enumeration, the field representative uses the control card to record a few basic characteristics of the housing unit, including the telephone number which is recorded for use in callbacks, reinterviews, and subsequent enumerations. Basically, the AHS-National and AHS-MS questions are the same. Separate questionnaires are used for occupied and unoccupied housing units, forms AHS-22 and AHS-23 for the national survey and forms AHS-62 and AHS-63 for the metropolitan survey, respectively. The questionnaire for occupied units has two parts, one for regular occupied and the other for usual residence elsewhere (URE) units (for example, seasonal units).

During the first decade (1973-1983) of the AHS, the overall design, purpose, and methodology did not change. This was particularly important, because the AHS is both a cross-sectional and a longitudinal survey. To derive the most benefits from the longitudinal data, year-to-year consistency in the way the Census Bureau conducted the survey was highly desirable. As different priorities emerged, HUD added new questions or subjects, such as neighborhood quality and energy consumption. The AHS questionnaire was revised for the redesigned sample in 1985 and is undergoing another revision for 1997.

Items Added

The redesigned questionnaire contains more data items than in previous years. Certain topics that had been included in prior years are now addressed in more detail. These expanded topics include: housing costs in general, mortgage information, fuels used for purposes other than heating, neighborhood land use, and information on the physical condition of the sample units. New topics added to the survey include: unit size (in square feet), lot size, the presence and age of major appliances, and other information on physical aspects of the unit (such as foundation type and presence of fireplaces and porches). The AHS now collects more information for vacant units, primarily concerned with the physical aspects and condition of the place, but it also tries to identify time-shared and vacation homes.

Items Dropped

Not many questions were dropped in the redesign. However, most of the information on the physical description of recent movers’ previous residences was cut, along with selected questions on the sample unit’s condition, and most of the income questions for nonrelative household members.

Questionnaire Content

A brief description of the 1985 AHS-National questionnaire content follows.

Information about the household. Number of persons; their age, sex, education, race, date moved into unit, and other demographic information; amounts and sources of income.

Information about the unit. Tenure, number and type of rooms, size of unit in square feet, number of units in the structure, year purchased, year built, type of basement, and other characteristics.

Equipment and facilities. Type of main and supplemental heating equipment; presence of appliances such as washer, dryer, air conditioner, garbage disposal, dishwasher; source of water; type of sewage disposal; type of parking facilities; etc.

Housing costs. Mortgage costs; real estate taxes; condominium fees; rent; utility costs; homeowner’s/household insurance; mobile home park fees; cost of repairs, alterations, and additions to the unit; cost of routine maintenance on the unit.

QUALITY INDICATORS

The building. Water leakage; blown fuses; water interruptions; toilet breakdowns; sewage breakdowns; peeling paint; broken plaster; holes in the floors, walls, or ceilings; signs of rats, etc.

The neighborhood. Trash in neighborhood, vandalized buildings, condition of streets.

Recent mover information. Reason for move, why this unit/neighborhood chosen, location of previous unit, tenure, and household size of previous residence, and other data.

Journey-to-work. Miles travelled to work, type of transportation used, location of job.

In addition to the core questionnaire, sometimes add-on questions or supplements are used. Contents of some supplements are as follows:

Mobility supplement. Location of birthplace, location of residence when 16, likelihood of moving from current residence within the next 5 years, preference for location of residence in 5 years.

Neighborhood quality supplement. Presence of conditions such as street noise, trash, crime, or commercial establishments in neighborhood; quality of local police, hospitals, public transportation, shopping, and schools.

Components of inventory change supplement. Determines status of sample units, whether they are newly created units (new construction, house/mobile home moved in, converted from nonresidential use, or the result of a conversion to more units) or are returning units. It also determines the disposition of the previous inventory (demolition, disaster loss, merger with another unit, or some other loss from the inventory).

QUESTIONNAIRE RESEARCH AND DEVELOPMENT

It is well-known that the questionnaire design; for example, wording of questions and order in which questions and possible response categories for a question are presented, affects responses. The 1985 questionnaire for AHS-National was finalized after receiving comments from regional offices and pretesting new questions in trial interviews to minimize response errors. As part of a continuing effort to improve the questionnaire, field representatives were requested to evaluate 1988 AHS-MS questionnaires (AHS-62 and AHS-63) and describe any problems on an evaluation sheet after they had completed the field work. Hayes (1989) recorded comments made by field representatives. These comments indicated that respondents had problems in understanding some questions. Need for better classification of buildings, basements, toilet breakdown, sewage breakdown, public/private water system, etc. were indicated.

Studies of the ''reason-for-move'' question over the years (see Montfort (1983a), Montfort (1983b), and Masumura (1981)) provide an interesting example of the development of a question over time and its impact on data quality. The ''reason-for-move'' question investigates two topics. The first is the reason that an individual (the reference person) moved away from his/her last place of residence, and the second is the reason the individual chose his/her present residence. The ''reason-for-move-from'' has been a part of the AHS from the beginning. During the first years, the survey inquired about the *main* ''reason-for-move-from,'' then in 1978 the respondent was asked for *all* reasons for the move, and in the followup question he/she was asked to choose the *main* ''reason-for-move-from.'' In 1979, the ''reason-for-move-to'' question was added to the survey. These questions are still part of the survey although much rewording has been done over the years. Essentially the format is the same, the question asks for *all* reasons and follows it up by asking the respondent to pick the *main* reason. There was a ''recent movers'' supplement in 1985 which included versions of these items.

There seems to be many problems with this question, or at least aspects of this question, which make it difficult to analyze. First, the question has many categories as possible responses and among those categories is the ''other'' response. The ''other'' response is chosen quite frequently, which means that there are many write-in answers and suggests that the setup of the categories is not ideal. This also is evidenced by the fact that the question has gone through so much rewording over the years. Second, there is, as always, a problem following the procedures for these questions. A common mistake being made is when more than one reason is given for the *main* reason. Finally, the similarity of the ''reason-for-move-to'' and the ''reason-for-move-from'' questions is possibly a source of confusion. It may not be clear whether the respondent is differentiating between these two topics, although it should be pointed out that the parallel wording between the ''reason-for-move-to'' and the ''reason-for-move-from'' questions was greatly diminished in the 1985 AHS-National.

INTERVIEW TIME AND COST

The length of a household interview depends in part on whether the housing unit is occupied or vacant. After the introduction of the redesigned sample and the new questionnaire in 1985, it was observed that field representatives were taking more time to complete interviews when compared to the 1983 AHS-MS survey. A time-study was conducted during the 1985 October (panel 10) enumeration for the AHS-MS to estimate time spent per interview for

regular occupied, usual residence elsewhere (URE), and vacant units. The results obtained from this study are summarized below from Quansey (1986).

Number of visits to obtain the initial interview. The total number of visits made to the sample unit in order to obtain the initial interview was recorded. For the 11 metropolitan areas, the average number of visits for regular interviews was 2.77 visits (ranging from a low of 2.21 visits for the Tampa-St. Petersburg MSA to a high of 3.01 visits for the Dallas PMSA). The average number of visits for the 22 URE cases was 3.55, with considerable variation around this figure due to the low number of cases (from an average of 1 visit for the 2 URE cases in the Detroit PMSA to a high average of 7.50 visits for the 2 URE cases in the San Francisco-Oakland PMSA's). The average number of visits for vacant interviews ranged from 2.05 for the Detroit PMSA to 3.23 for the Washington, DC MSA, with an average of 2.56 visits for all metropolitan areas. Noteworthy here are the multiple number of visits made to each unit in order to obtain an interview, with an overall average of 2.76 visits per unit (irrespective of the type of interview).

In-house interviewing time. Clocking in-house interviewing time begins once a person answers the door. Interviewing time spans the time it takes to complete the control card and appropriate questionnaire items. For the 11 metropolitan areas, average in-house interviewing time was 37.01 minutes for regular interviews, 22.05 minutes for URE interviews, and 19.38 minutes for vacant interviews. The range for regular interviews was from 34.38 minutes (Minneapolis-St. Paul MSA) to 42.51 minutes (Washington, DC MSA). URE interviews ranged from 16.50 minutes (Dallas PMSA) to 30 minutes (San Francisco-Oakland PMSA's). Average in-house interviewing time for vacant interviews ranged from 14.93 minutes (Boston CMSA) to 25.62 minutes (San Francisco-Oakland PMSA's).

Time for unit size measurement. A unit size measurement was not required if the respondent provided unit size information during the interview. If the respondent could not provide unit size, a measurement was performed for 1-unit-detached buildings and mobile homes. The average number of minutes required to perform a unit measurement for regular interviews ranged from 4.63 (Los Angeles-Long Beach PMSA) to 10.17 (Fort Worth-Arlington PMSA), for an overall average of 7.16 minutes per case. The one URE interview with measurement data had recorded 5 minutes. The average time for vacant interviews ranged from 2 minutes (Philadelphia PMSA) to 15 minutes (Detroit PMSA), and an overall average for all metropolitan areas was 8.64 minutes per case.

Field representative observation items. Time spent completing the field representative observation items in the questionnaire, regarding the external physical characteristics of the sample housing unit and the area within 300 feet, also was recorded separately. The average number of minutes spent completing these items showed little variation by type of interview: 2.81 minutes for regular interviews, 2.50 minutes for URE interviews, and 2.86 minutes for vacant interviews. The range for regular interviews was from 1.69 minutes (Minneapolis-St. Paul MSA) to 4.61 minutes (Fort Worth- Arlington PMSA). For URE interviews, the average time for completing observation items ranged from 1 minute (Detroit and Dallas PMSA's) to 4 minutes (Washington DC MSA and Fort Worth-Arlington PMSA). The range for vacant interviews was from a low of 1.63 minutes (Minneapolis- St. Paul MSA) to a high of 5.38 minutes (San Francisco- Oakland PMSA).

Callbacks. Subsequent to the initial interview, field representatives sometimes need to make callbacks by telephone or in person to obtain mortgage, mobility, income, or other missing information, or to clarify a given response. Because the callback portion of the time-study was designed to estimate the total time the callback effort added to the length of the interview, even unsuccessful attempts to reach a respondent by telephone were to be recorded. Of the 6,460 interviews reported during this time-study, only 794 (12.3 percent) contained any callback information. Field representatives were instructed to mark a ''Not applicable'' box if no callbacks were made for a unit, but for the vast majority of cases the entire item was left blank. It is impossible to ascertain if only 12.3 percent of the cases required callbacks, or if the entire section (printed on the back of the form AHS-60) had been overlooked. Because this item may be seriously underreported, detailed analysis of the data should be done with great caution.

For the 11 metropolitan areas, 86.5 percent of the callbacks were made by telephone. The predominant reason stated for the callbacks (69.9 percent) was to obtain mobility information. Fewer than 1 percent of the callbacks were made to obtain information for something other than mobility, mortgage, or income items. The average number of callbacks made per unit was 1.63, with very little dispersion among individual metropolitan areas. Each callback lasted, on the average, 3.10 minutes, with a low average of 1.83 minutes for Boston CMSA and a high average of 6.33 minutes for Phoenix MSA. For all metropolitan areas, the average time per unit spent obtaining callback information was 5.07 minutes, ranging from 2.38 minutes (Boston CMSA) to 7.39 minutes (Phoenix MSA).

Time per interview. Table 3.1 presents average time spent per unit by interview type for each metropolitan area. The average time per unit is derived from the combination of all four time components (in-house interviewing time, unit measurement, observation items, callbacks). The time for a regular interview ranges from a low average of 35.44 minutes (Boston CMSA) to a high average of 46.68 minutes (Tampa-St. Petersburg MSA); and the average for all metropolitan areas is 41.17 minutes. The average time for a URE interview is 24.67 minutes for all metropolitan areas,

Table 3.1. **Average Time per Interview for AHS-MS, 1985**

(In minutes)

Metropolitan areas	Housing unit		
	Regular	URE	Vacant
All	**41.17**	**24.67**	**22.52**
Boston, MA-NH, CMSA	35.44	25.00	18.70
Philadelphia, PA-NJ, PMSA	41.98	-	24.67
Detroit, MI, PMSA	37.94	23.50	20.85
Minneapolis-St. Paul, MN-WI, MSA	37.37	22.00	19.10
San Francisco-Oakland, CA, PMSA	45.39	32.00	31.96
Washington, DC-MD-VA, MSA	39.40	28.00	21.91
Tampa-St. Petersburg, FL, MSA	46.68	25.00	25.22
Dallas, TX, PMSA	40.96	17.50	22.25
Fort Worth-Arlington, TX, PMSA	41.79	21.75	19.47
Phoenix, AZ, MSA	39.40	-	18.46
Los Angeles-Long Beach, CA, PMSA	45.54	27.60	28.03

- Not available.

Source: Quansey (1986).

ranging from a low average of 17.5 minutes (Dallas PMSA) to a high average of 27.6 minutes (Los Angeles-Long Beach PMSA). For vacant interviews, the average time is 22.52 minutes for all metropolitan areas, ranging from 18.7 minutes (Boston CMSA) to 31.96 minutes (San Francisco-Oakland PMSA's). The overall average time for all three types of interviews and all metropolitan areas is 39.70 minutes per interview.

Results of this time-study effort represent self-reporting by field representatives during one panel which comprised the seventh interviewing month for the 1985 AHS-MS. Figures compiled for unit measurement and callback information are derived from a much smaller base than used to average the other time-study components. Estimates of ''Out-of-house'' factors contributing to the number of minutes per case, such as time devoted to planning an itinerary, listing prior to interviewing, travel time, editing time, and time spent filling out payroll forms, cannot be derived from this study. Because of this, actual interviewing ''Minutes Per Unit'' during 1985 AHS-MS was higher.

No direct estimate of time spent in personal visit (PV) interview for AHS-National is available, but except for travel time, other components of interview time are likely to be similar. However, the time for telephone interviews, both computer assisted telephone interview (CATI) and non-CATI, for AHS-National may be different.

POTENTIAL SOURCES OF ERRORS IN THE DATA COLLECTION PROCEDURE

The potential sources of nonsampling error in the AHS data collection procedure are many; for example, listing error, nonresponse, simple and correlated response variance, interview mode, questionnaire, problems with year built, problems with multiunit structures, etc. Some of these errors are systematically investigated and controlled as part of the AHS reinterview program. In this section we discuss some sources of errors. Others, for example, nonresponse and measurement errors, are discussed in chapters four and five.

Listing by Observation in Area Segments

Field representatives try to obtain address information by observation. Inquiry at a housing unit is made only when necessary. Most of the listing errors occur in area segments (Schreiner, 1977). In the fall of 1975, a rural listing test was conducted in nine counties in the South (Louisiana, Mississippi, and Arkansas) to investigate the feasibility of conducting the census in rural areas by mail. The use of a ''knock on every door'' procedure generally achieved statistically significant coverage improvement over the procedure used by the Census Bureau in CPS, AHS, and other household surveys, but its cost could be prohibitive and could result in undue respondent burden. A modified ''knock on every door'' procedure that allowed a simple call back as a last resort appeared to obtain enough additional coverage to offset the increased cost (Dinwiddie, 1977).

An intensive coverage check was done for CPS in October 1966 and June 1967. This check was to identify units missed because they were not listed or interviewed and to identify units incorrectly included in the sample. The results summarized by Shapiro (1980) showed that there was a net undercoverage of about 1.9 percent in area segments. This situation has improved in recent years with better maps and procedures. For example, in 1988, a net error rate of -0.85 percent (S.E. = 0.41 percent) in area segments in the CPS reveals a slight undercount of units in the original listing (Waite, 1990c). An evaluation of listing errors for AHS is not available but their magnitude is likely to be similar to those of CPS.

Problems With the Coverage Improvement Screening Procedure

In coverage improvement segments, a screening procedure (form AHS-215) is used to determine the type of living quarter (mobile home or nonmobile home), permit-issuing area or nonpermit-issuing area, when the structure was built and whether this structure contained any living quarters on April 1, 1980 (Matchett, 1985, 1987). The procedure does not always perform its intended function because the year built is misreported by some respondents. It is often difficult for a respondent to determine the year a structure was built, particularly if he/she is not the first owner or if he/she is a renter. As a result, some units are incorrectly omitted from the coverage improvement sample while some other units are incorrectly included. The extent of coverage error due to misreporting of year built is not known. The response error in the year built data is discussed in chapter 5.

Respondent Rule and Its Effect on Data.

Exhibit 3.1 provides the respondent rules for AHS-National as given in the 1993 AHS-National Field Representative Manual. Respondent rules for AHS-MS are similar.

The potential sources of errors for the respondent selected by these rules are respondents not knowing answers, not willing to provide them, or providing faulty answers. Reinterviews are conducted to determine the extent of this problem. Discrepancies in responses are discussed in the section, ''Response Errors,'' chapter 5, page 48, in connection with reinterview results.

Field representatives are not allowed to use a proxy response for a regular interview if no knowledgeable household member 16 years of age or older is available without contacting the regional office.
Wells (1982) documents the results of a review of proxy interviews for the 1981 AHS-National by Washington staff, which shows few cases of proxy interviews that were classified as (Type-A-05) noninterview due to ''ineligible respondent'' noninterview. This shows that the field representatives follow the rules for the selection of a respondent correctly in most cases.

NONINTERVIEWS

A noninterview occurs when a field representative cannot obtain an interview for a housing unit, occupied or vacant, that is eligible for interview. A noninterview for an occupied housing unit—regular occupied or usual residence elsewhere (URE) unit—is called Type A noninterview.

The reasons for Type A noninterview as given in the questionnaire (AHS-22) for national and (AHS-62) for metropolitan occupied units are:

Type A

01	☐	No one home
02	☐	Temporarily absent
03	☐	Refused
04	☐	Unable to locate
05	☐	Other occupied Specify_____

The ''no-one-home'' households are those whose members cannot be contacted at home by the field representatives after repeated calls. ''Temporarily absent'' households are those whose members are away on vacation, business trips, etc. and will not be available for interview during the survey period. ''Refusal'' households are those which are contacted but whose members refuse to respond. ''Unable-to-locate'' housing units are those which field representatives cannot physically locate.

Noninterviews for unoccupied units have been classified as Type B or Type C. Units which are not eligible for interview at the present time, but could become eligible in the future are Type B noninterviews. Units ineligible for sample, either because they no longer exist or because of sampling reasons, are Type C noninterviews.

Reasons for Type B or Type C noninterviews as given in the questionnaire (AHS-23) for national and (AHS-63) for metropolitan unoccupied units are as follows:

Type B

10	☐	Permit granted, construction not started
11	☐	Under construction, not ready
12	☐	Permanent or temporary business or commercial storage
13	☐	Unoccupied site for mobile home or tent
14	☐	OTHER unit or converted to nonstaff
15	☐	Occupancy prohibited
16	☐	Interior exposed to the elements
17	☐	Type B, not classified above Specify_____
19	☐	Updating code 2, 3, 5, or 11 (codes 2, 3, 5, and 11 refer to change in status of units from the previous enumeration)

Type C

30	☐	Demolished or disaster loss
31	☐	House or mobile home moved
33	☐	Merged not in current sample
36	☐	Permit abandoned
37	☐	Type C, not classified above Specify_____

The number of interviewed housing units and number of Type A, Type B, and Type C noninterviews for AHS-National for 1985, 1987, 1989, 1991, and 1993 are given in table 3.2.It can be seen that the designated sample size (that is, number of housing units) was 53,895 but the effective sample size (that is, number of interviewed units) was 48,830 in the 1985 survey.The sample size for AHS-National increased in 1987 and 1989 largely due to the addition of new construction units.In addition to the basic sample, there were two supplemental samples (1) neighborhood sample in 1985, 1989,and 1993, and (2) rural sample in 1987 and 1991.The sample size data in table 3.2 include these supplemental samples.The Type A noninterview rates were 4.2 percent, 3.2 percent, 4.2 percent, 4.4 percent, and 4.1 percent in 1985, 1987, 1989, 1991, and 1993.The type B and Type C noninterviews are out of scope at the time of the survey, if there are no errors due to misclassification by field representatives.Type B and Type C noninterviews are recorded to keep track of designated housing units and to correct misclassifications.Errors in classification of Type B and Type C noninterviews are discussed in the next section.

Exhibit 3.1. **Respondent Rules for AHS-National**

Who is an eligible respondent

1. Regular interview	For a *regular* interview, *any knowledgeable adult household member ("1" circled in Control Card item 14) 16 years of age or older is technically eligible* to act as the respondent. However, try to interview the most knowledgeable household member; that is, one who appears to know—or might reasonably be expected to know—the answers to all or the majority of the questions.This will frequently be the reference person or his/her spouse.
Knowledgeable household member unavailable	If no knowledgeable household member 16 years of age or older is available, try to determine from some reliable source when an eligible respondent will be available for interview. *Do not use a proxy respondent.* If an unusual situation exists, contact your regional office.
2. URE Interview	In a unit occupied *entirely* by persons with usual residence elsewhere, conduct a personal interview with the most knowledgeable occupant 16 years of age or older.
3. Vacant interview	If a unit meets the definition of a *vacant* interview, interview the owner, agent, or resident or building manager.Consider a janitor as an agent if he/she is responsible for answering inquires about the unit. (This person does *not* have to be interviewed at the sample unit.) Frequently, the name, address, and phone number of persons who can provide information are posted on the property. Interview a knowledgeable neighbor *only* when the landlord, owner, or agent cannot be interviewed. If a neighbor supplies some of the information but refers you to the owner for the rest, interview the owner for all items in the questionnaire including those items for which the answers were supplied by the neighbor. If the owner or agent is outside your assignment area, and you cannot obtain the necessary information from another acceptable source, contact the office. The RO will transfer the case to get an interview in person from the knowledgeable respondent.
4. Language difficulties	If the occupants of the sample unit do not understand or speak English, you may conduct the interview *through* an interpreter.The interpreter must translate the questions, *not* answer them from personal knowledge or observation. When conducting an interview in which an interpreter is needed because of a language problem, ask the respondent if he/she is willing to have another person act as interpreter. If the respondent objects or you cannot locate an interpreter nearby at the time of the interview, call the office to see if another field representative who speaks the respondent's language can conduct the interview later. When an interpreter is used, a Form BC-1415, Contract for Interpreter Service, must be completed.Reimbursement may not be made if the BC-1415 is not sent in with your other payroll forms. Refer to your 11-55 Administrative Handbook for more detailed information on the use of interpreters. The person providing the information to the interpreter must qualify as an eligible respondent as defined above. The interpreter could be someone such as a family member, a neighbor, an official interpreter, or even you, if you speak that language. When an interview is conducted *through* an interpreter, the occupant answering the questions is considered to be the respondent. If you are unable to obtain an interpreter, contact your regional office for instructions.

Table 3.2. **Number of Interviews and Noninterviews by Type for AHS-National for 1985-1993**

	1985		1987		1989		1991		1993	
	Number of units	Percent	Number of units	Percent	Number of units	Percent	Number of units	Percent	Number of units	Percent
Total interviews	**48,830**	**90.7**	**49,641**	**91.0**	**51,823**	**88.0**	**51,027**	**86.0**	**55,981**	**86.2**
Regular interviews	**43,104**	**80.0**	**43,436**	**79.6**	**45,772**	**77.7**	**44,764**	**75.2**	**49,326**	**75.9**
URE interviews	**469**	**0.9**	**481**	**0.9**	**572**	**1.0**	**594**	**1.0**	**641**	**1.0**
Vacant interviews	**5,257**	**9.8**	**5,724**	**10.5**	**5,479**	**9.3**	**5,669**	**9.5**	**6,014**	**9.3**
Type A noninterviews										
Codes										
01	272	0.5	224	0.4	251	0.4	196	0.3	120	0.2
02	56	0.1	45	0.1	36	0.1	51	0.1	47	0.1
03	1313	2.4	1174	2.2	1715	2.9	1,917	3.2	2,083	3.2
04	163	0.3	54	0.1	17	0.0	19	0.0	1	0.0
05	361	0.7	132	0.2	250	0.4	142	0.2	161	0.2
Total Type A	**2,165**	**4.0 [1](4.2)**	**1,629**	**3.0 [1](3.2)**	**2,269**	**3.8 [1](4.2)**	**2,325**	**3.8 [1](4.4)**	**2,412**	**3.7 [1](4.1)**
Type B noninterviews										
Codes										
10	34	0.1	43	0.1	39	0.1	59	0.1	33	0.1
11	103	0.2	140	0.3	97	0.2	66	0.1	64	0.1
12	266	0.5	352	0.6	375	0.6	472	0.8	440	0.7
13	187	0.3	256	0.5	214	0.4	261	0.4	252	0.4
14	372	0.7	404	0.7	441	0.7	467	0.8	466	0.7
15	50	0.1	78	0.1	90	0.2	82	0.1	111	0.2
16	278	0.5	357	0.7	333	0.6	374	0.6	340	0.5
17	58	0.1	58	0.1	82	0.1	80	0.1	97	0.1
19	337	0.6	505	0.9	531	0.9	681	1.1	700	1.1
Total Type B	**1,685**	**3.1**	**2,193**	**4.0**	**2,202**	**3.7**	**2,542**	**4.1**	**2,503**	**3.9**
Type C noninterviews										
Codes										
30	585	1.1	427	0.8	1238	2.1	1,665	2.8	1,821	2.8
31	334	0.6	313	0.6	656	1.1	971	1.6	919	1.4
33	18	0.0	5	0.0	38	0.1	35	0.1	65	0.1
36	54	0.1	38	0.1	119	0.2	148	0.2	155	0.2
37	224	0.4	311	0.6	597	1.0	778	1.3	1,142	1.8
Total Type C	**1,215**	**2.2**	**1,094**	**2.0**	**2,648**	**4.5**	**3,597**	**6.0**	**4,102**	**6.3**
Designated sample size	53,895	100.0	54,557	100.0	58,942	100.0	59,491	100.0	64,998	100.0

[1] "Official" Type A noninterview rates based on Type A noninterview/(interview plus Type A noninterview).

The number of interviewed housing units and Type A noninterview rates for all metropolitan areas in the AHS-MS for 1986-1994 are given in table 3.3. Type A noninterview rates varied by metropolitan area and by year.

In 1986, Type A noninterview rate ranged from a low of 2.3 percent in Cincinnati to a high of 6.5 percent in San Antonio. Only two metropolitan areas had a noninterview rate above 5 percent. In 1987, Type A noninterview rate exceeded 5 percent in four metropolitan areas. In 1988, Type A noninterview rate was below 5 percent in all 11 metropolitan areas. In 1989, Type A noninterview rate ranged from a low of 2.4 percent in Minneapolis to a high of 9.1 percent in Washington, DC. The noninterview rate exceeded 5 percent in 7 out of 11 metropolitan areas. In 1990, Type A noninterview rate ranged from a low of 2.3 percent in Cincinnati to a high of 8.7 percent in Anaheim. The noninterview rate exceeded 5 percent in 3 out 11 metropolitan areas in 1990. In 1991, Type A noninterview rate ranged from a low of 2.7 percent in St. Louis to a high of 8.5 percent in Northern New Jersey. The noninterview rate exceeded 5 percent in 2 out of 11 metropolitan areas in 1991. In 1992, Type A noninterview rate ranged from a low of 2.9 percent in Birmingham to a high of 5.6 percent in Norfolk. The noninterview rate exceeded 5 percent in 1 out of 8 metropolitan areas in 1992. In 1993, Type A noninterview rate ranged from a low of 4.6 percent in Detroit to a high of 8.4 percent in Washington, DC. The noninterview rate exceeded 5 percent in 3 out of 7 metropolitan areas in 1993. In 1994, Type A noninterview rate ranged from a low of 2.6 percent in San Diego to a high of 6.1 percent in Anaheim. The noninterview rate exceeded 5 percent in 3 out of 8 metropolitan areas in 1994 (see table 3.3).

Table 3.3. **Number of Interviews and Type A Noninterviews for AHS-MS 1986-1994**

Metropolitan areas	Number of interviews	Type A non-interview (percent)
1986		
Anaheim-Santa Ana, CA PMSA	3,047	5.8
Cincinnati, OH-KY-IN PMSA	3,002	2.3
Denver, CO CMSA	2,938	4.7
Kansas City. MO-KS CMSA	3,072	3.1
Miami-Ft. Lauderdale, FL CMSA	2,877	4.5
New Orleans, LA MSA	2,943	4.3
Pittsburgh, PA CMSA	2,842	3.1
Portland, OR-WA CMSA	2,976	3.3
Riverside-San Bernardino-Ontario, CA PMSA	2,934	3.0
Rochester, NY MSA	2,982	4.3
San Antonio, TX MSA	2,928	6.5
1987		
Atlanta, GA MSA	3,474	8.7
Baltimore, MD MSA	3,404	4.4
Chicago, IL area PMSA's	4,304	5.6
Columbus, OH MSA	3,244	4.3
Hartford, CT CMSA	3,315	4.2
Houston, TX area PMSA's	3,413	4.0
New York-Nassau-Suffolk, NY area PMSA's	4,972	6.8
Northern NJ area PMSA's	3,926	8.4
St. Louis, MO-IL CMSA	3,382	3.9
San Diego, CA MSA	3,392	3.4
Seattle-Tacoma, WA CMSA	3,335	4.6
1988		
Birmingham, AL MSA	3,272	4.2
Buffalo, NY CMSA	3,466	4.4
Cleveland, OH PMSA	3,417	4.7
Indianapolis, IN MSA	3,592	2.4
Memphis, TN-AR-MS MSA	3,775	3.5
Milwaukee, WI PMSA	3,586	2.9
Norfolk-Virginia Beach-Newport News, VA MSA	3,978	3.9
Oklahoma City, OK MSA	3,520	4.7
Providence-Pawtucket-Warwick, RI-MA PMSA's	3,776	3.7
Salt Lake City, UT MSA	3,752	3.9
San Jose, CA PMSA	3,743	4.0
1989		
Boston, MA-NH CMSA	4,000	5.1
Dallas, TX PMSA	3,520	5.4
Detroit, MI PMSA	3,723	5.1
Fort Worth-Arlington, TX PMSA	3,295	4.5
Los Angeles-Long Beach, CA PMSA	4,438	8.2
Minneapolis-St Paul, MN-WI PMSA	3,780	2.4
Philadelphia, PA-NJ PMSA	3,835	6.6
Phoenix, AZ MSA	3,755	4.1
San Francisco-Oakland, CA area PMSA's	3,866	6.6
Tampa-St. Petersburg, FL MSA	3,699	4.1
Washington, DC-MD-VA MSA	3,789	9.1
1990		
Anaheim-Santa Ana, CA PMSA	4,343	8.7
Cincinnati, OH-KY-IN PMSA	4,156	2.3
Denver, CO CMSA	4,200	5.2
Kansas City, MO-KS CMSA	4,237	3.5
Miami-Ft. Lauderdale, FL CMSA	4,684	3.5
New Orleans, LA MSA	3,836	3.7
Pittsburgh, PA CMSA	3,704	4.0
Portland, OR-WA CMSA	4,300	3.2
Riverside-San Bernardino-Ontario, CA PMSA	4,880	7.2
Rochester, NY MSA	4,188	2.6
San Antonio, TX, MSA	4,108	2.8
1991		
Atlanta, GA MSA	4,364	4.6
Baltimore, MD MSA	4,074	5.0
Chicago, IL area PMSA's	5,066	3.2
Columbus, OH MSA	3,965	4.4
Hartford, CT CMSA	3,979	3.3
Houston, TX area PMSA's	3,782	3.9
New York-Nassau-Suffolk, NY area PMSA's	5,770	8.1
Northern NJ area PMSA's	4,844	8.5
St. Louis, MO-IL CMSA	4,041	2.7
San Diego, CA MSA	4,170	4.1
Seattle-Tacoma, WA CMSA	4,134	4.7
1992		
Birmingham, AL MSA	3,882	2.9
Cleveland, OH PMSA	3,906	4.1
Indianapolis, IN MSA	4,223	2.9
Memphis, TN-AR-MS MSA	4,468	3.0
Norfolk-Virginia Beach-Newport News, VA MSA	4,678	5.6
Oklahoma City, OK MSA	4,006	4.1
Providence-Pawtucket-Warwick, RI-MA PMSA's	4,424	3.3
Salt Like City, UT MSA	4,343	3.2
1993		
Boston, MA-NH CMSA	4,348	4.7
Detroit, MI PMSA	4,024	4.6
Minneapolis-St Paul, MN-WI PMSA	4,353	5.8
San Francisco-Oakland, CA area PMSA's	4,314	6.7
San Jose, CA PMSA	4,294	4.9
Tampa-St. Petersburg, FL MSA	4,280	4.6
Washington, DC-MD-VA MSA	4,516	8.4
1994		
Anaheim-Santa Ana, CA PMSA	3,846	6.1
Buffalo, NY MSA	3,659	3.9
Dallas, TX PMSA	3,696	5.8
Fort Worth-Arlington, TX PMSA	3,441	4.7
Milwaukee, WI PMSA	3,712	4.5
Phoenix, AZ MSA	4,150	5.7
Riverside-San Bernardino-Ontario, CA PMSA	4,489	3.3
San Diego, CA MSA	3,854	2.6

THE ''BUILDING LOSS—VACANT OTHER'' RECHECK PROGRAM FOR AHS-NATIONAL

The ''Building Loss—Vacant Other'' recheck program refers to verifying the classification of a housing unit. This program was conducted for AHS-National. For these studies the classification of all noninterview housing units (Type B and Type C) and all ''Vacant-Other'' units was checked. Also, the coding of the ''type of living quarters'' for the housing unit was checked. Once a unit is classified as a Type C noninterview, the building is considered a permanent loss and the unit will not be surveyed in the future. On the other hand, a unit classified as a Type B noninterview may be eligible for interview in the future. The importance of this problem is obvious. Buildings that are classified as losses when they are not losses or vice versa will bias the results of the survey. The more misclassifications there are, the more serious this problem will become.

The ''Building Loss—Vacant Other'' recheck program was first conducted in the summer of 1974 to check and correct misclassifications. The program was a review of units classified in the Type B and Type C noninterview loss categories in 1973 and 1974, as well as units classified as ''Vacant-Other'' in 1974. Before the 1977 National survey, only Type B and Type C units not classified the previous year as a Type B or Type C and units classified as Vacant-Other were selected for recheck. These cases were chosen because it was felt they were more error-prone than units with an established history as a noninterview or vacant. By contrast, since 1977 the building loss review also has included a recheck program for Type B and Type C noninterview units not classified as the same Type B or Type C code the previous year and units classified as Vacant-Other. For these, there was a review of all pertinent information (control cards, questionnaires for vacant units, field representative comments, and reinterview forms AHS-393's for all reported Type B and Type C noninterviews).

Starting in 1978, units confirmed as Vacant-Other were not reinterviewed for the current survey year if the unit was classified again as Vacant-Other. To qualify as a confirmed Vacant-Other, the unit must have been Vacant-Other the 2 years prior to the current enumeration and have a duration of vacancy of 12 months or more during the current enumeration. In 1979, the AHS-397 checklist for Type B and Type C noninterviews was introduced.

The information sources for the 1978 and 1979 program included: questionnaires (AHS-23); control cards (AHS-1), Inter-Comms (11-36), Reinterview forms (AHS-393) where available, records from the 1978 Building Loss—Vacant Other recheck (1979 only), AHS-397 Checklist for Type B and C noninterviews (1979 only), the resources of the Regional Offices (such as listing sheets and field representatives' statements), and Demographic Statistical Methods Division (DSMD) research (checking sample reports, listing sheets, etc.). The AHS-393 is filled in for first time Type B and Type C noninterviews and any Type B that changes categories. The AHS-397 is filled in for all Type B and Type C noninterviews. This form helps guide the field representative to the appropriate classification and provides HHES with more detailed information about the unit and the explanation of what led the field representative to the original classification of these units. If the reinterview contradicted the original field representative's classification, but after reviewing all available sources of information there was still a question as to which was the correct classification, the original classification was accepted.

Misclassifications were corrected as they were discovered. In certain cases, the unit was coded as a Type A noninterview, even though the correct classification could not be determined, because of processing requirements. (If the case is reclassified as a vacant or occupied interview, the information recorded on the control card and questionnaire in noninterview cases is not sufficient for the questionnaire to pass the incomplete document check.) These units were classified a Type A-5, ''other-occupied,'' and marked as either occupied or vacant in the occupancy status item. These were later adjusted for during the application of the noninterview adjustment factor during the weighting.

Detailed data from this program over the years are provided in two documents by Williams (1978, 1979). In this report, a summary of important results is provided from Chakrabarty (1992b). Table 3.4 provides misclassification error by category for Type B noninterviews in 1978 and 1979. The overall error rate for Type B was 7.1 percent in 1978 and 8.3 percent in 1979. The error rates by categories varied considerably, but rates based on small numbers of cases; for example, 35 cases for the ''scheduled to be demolished'' category, are not reliable. However, error rates for ''scheduled to be demolished'' and ''other'' are much higher than rates for other categories. Table 3.5 provides misclassification error by category for Type C noninterviews in 1978 and 1979. The overall error rates for Type C were 9.8 percent in 1978 and 11.6 percent in 1979. As in Type B, error rates by categories of Type C varied considerably.

In 1979, the recheck program increased the AHS sample size by 137 units by including 275 units that were incorrectly deleted and dropping 138 units that were incorrectly retained. Thus, it appears that at least some field representatives seem to have a tendency to delete units that should not be deleted (classify true Type B's as Type C's) rather than include units that should not be included (classify true Type C's as Type B's), resulting in a decrease in the size of the sample.

The coding of the ''type of living quarter'' for the housing unit also was checked. The results are provided in table 3.6. The misclassification error rate for a code reflects the percent of cases that were originally classified with that code but should not have been. It can be seen that the overall error rate in coding of ''type of living quarter'' by field representatives was 8.6 percent in 1977 and 5.4 percent in 1979. One of the most interesting cases is the ''HU permanent in transient hotel, motel, etc.'' code. In this case,

the error rate went from 74.1 percent to 50.0 percent to 0.0 percent in just 3 years. This is probably not entirely due to an improvement in the survey procedures, but rather partially due to the variability that is inherent in a code with a small number of cases. Another explanation for the high error rates in the codes with few cases is that these codes represent the unusual situations, with which field representatives have the most difficulty dealing.

Many recommendations for changes to the recheck and survey procedures were made in the last few years that the recheck program has been performed. An overall improvement in noninterview and vacancy code identification seems to be necessary in order to reduce the number of misclassifications in all categories. The AHS-393 and AHS-397 checklists were a step in the right direction in helping field representatives to identify the proper codes for housing units. These were used during the 1979 recheck. Other modifications also have been implemented. More detailed instructions have been added to the field representatives' manual, and further changes in field representatives instructions would probably be helpful. The type of unit which frequently has given field representatives trouble is the ''Vacant-Other'' unit , which does not fit into any of the vacant codes. Some errors in the classification are to be expected. The recheck program identified problems, helped improve procedures, and improved the quality of AHS-National data.

The Census Bureau discontinued the building loss recheck program in 1987. Prior to that date, the scope of the program had been reduced over time so that only Type C noninterviews and Type B ''other'' noninterviews were included. The reasons for the cutback and eventual discontinuance of the program were chiefly related to cost. The building loss recheck program used large amounts of staff time. During the first several years of the program, the information gained was used to modify the training, manuals, and data collection forms for losses to clarify some of the problem areas. Therefore, as these improvements were incorporated, it was felt that such a large-scale review was not needed. The last recheck, in 1985, was used to determine if the redesigned questionnaire had helped field representatives classify losses. The results of the 1985 recheck were encouraging and the program was not reinstituted for the 1987 survey.

Table 3.4. **Misclassification Error for Type B Noninterviews for AHS-National**

Categories of Type B	Year					
	1978			1979		
	Number of cases	Misclassified		Number of cases	Misclassified	
		Number	Percent		Number	Percent
Unit for nonresidential use	787	37	(4.7)	814	48	(5.9)
Other unit, except unoccupied site for mobile home or tent	1,405	52	(3.7)	1,415	150	(10.6)
Unoccupied mobile home site	544	87	(16.0)	484	61	(12.6)
Under construction, not ready	500	9	(1.8)	385	5	(1.3)
Scheduled to be demolished	35	24	(68.6)	27	15	(55.6)
Condemned/unoccupied by law	73	20	(27.4)	82	4	(4.9)
Interior exposed to elements	704	38	(5.4)	758	25	(3.3)
Unit severely damaged by fire	59	13	(22.1)	43	3	(7.0)
Other	28	21	(75.0)	28	24	(85.7)
Permit granted, construction not started	90	0	(0.0)	102	6	(5.9)
Overall error rate	4,225	301	(7.1)	4,125	341	(8.3)

Source: Chakrabarty (1992b).

Table 3.5. **Misclassification Error for Type C Noninterviews for AHS-National**

Categories of Type C	Year					
	1978			1979		
	Number of cases	Misclassified		Number of cases	Misclassified	
		Number	Percent		Number	Percent
Unit eliminated in conversion	66	21	(31.8)	63	46	(73.0)
Demolished	247	24	(9.7)	233	37	(15.9)
Disaster loss (flood, etc.)	7	0	(0.0)	5	1	(20.0)
Disaster loss—fire	53	7	(13.2)	33	0	(0.0)
House or mobile home moved	457	48	(10.5)	434	33	(7.6)
Merged—not in current sample	100	10	(10.0)	80	6	(7.5)
Built after April 1, 1970	47	6	(12.8)	76	5	(6.6)
Other	618	42	(6.8)	425	31	(7.3)
Unused permit—abandoned	35	2	(5.7)	21	0	(0.0)
Overall error rate	1,630	160	(9.8)	1,370	159	(11.6)

Source: Chakrabarty (1992b).

Table 3.6. **Misclassification Error for Type of Living Quarter for AHS-National**

Type of living quarter	Year					
	1977		1978		1979	
	Misclassification		Misclassification		Misclassification	
	Error percent	Number of cases	Error percent	Number of cases	Error percent	Number of cases
House, apartment, flat	3.0	(2,900)	0.7	(2,867)	0.4	(2,993)
HU in nontransient hotel, etc.	0.0	(16)	0.0	(20)	0.0	(32)
HU permanent in transient hotel	74.1	(31)	50.0	(22)	0.0	(10)
HU in rooming house	3.5	(29)	3.8	(26)	0.0	(15)
Mobile home with no permanent room added	11.3	(177)	7.2	(207)	8.0	(237)
Mobile home with one or more permanent room added	11.1	(9)	0.0	(14)	10.5	(19)
HU not specified above	86.9	(46)	90.0	(30)	62.5	(16)
Quarters not HU in rooming or boarding house	33.3	(60)	43.4	(53)	22.6	(53)
HU not permanent in transient hotel, etc.	4.9	(265)	6.3	(207)	9.6	(94)
Unoccupied tent/trailer site	8.2	(524)	1.4	(443)	1.0	(404)
Other HU not specified above	15.4	(1,337)	11.1	(1,265)	9.7	(1,298)
Blank (for all Type C noninterview)	9.0	(2,000)	9.8	(1,630)	11.6	(1,370)
Overall error rate	8.6	(7,394)	6.2	(6,784)	5.4	(6,542)

Source: Chakrabarty (1992b).

Chapter 4.
Nonresponse Error

INTRODUCTION

In this report, ''nonresponse'' refers to two kinds of nonsampling error, noninterview and item nonresponse. A noninterview occurs when a field representative cannot obtain an interview for a housing unit, occupied or vacant. Reasons for noninterview are ''refusal,'' ''no-one-home,'' ''temporarily absent,'' or ''unable-to-locate.''

- Refusal households are those which are contacted but whose members refuse to respond.
- ''No-one-home'' households are those whose members cannot be contacted at home by the field representatives after repeated calls.
- ''Temporarily absent'' households are those whose members are away on vacation, business trips, etc. and will not be available for interview during the survey period.
- ''Unable-to-locate'' housing units are those which field representatives cannot physically locate.

In addition to these noninterviews, there may be one or more unanswered questions, referred to as item nonresponse, within interviewed units. Information on vacant units is collected from knowledgeable persons, neighbors, or landlords. Like occupied units, vacant units can have noninterview and item nonresponse. The type A noninterview for an occupied housing unit and type B or type C noninterview for a vacant unit are discussed in the section ''Noninterviews,'' chapter 3, page 29. Noninterview adjustments and imputation for item nonresponse are discussed in chapter 7.

STEPS TO MAXIMIZE RESPONSE RATES

Several steps are taken by the Census Bureau to encourage response to the AHS.

1. An advance letter from the Director of the Census Bureau explains the authority for and purposes of the survey, and urges participation. (see exhibit 4.1, page 39.)
2. Field representatives are trained to introduce themselves properly and to urge cooperation of respondents by explaining the purpose and importance of the survey. They carry official identification cards and portfolios identifying them as Census Bureau employees during personal visits. The ''ID'' card contains field representative's picture and signature.
3. If no one is home at the time of the first visit, field representatives determine the best time for a callback, either by asking neighbors or by telephoning later.
4. Field representatives assure respondents that their answers will be held in confidence and used only for statistical purposes.

UNABLE-TO-LOCATE UNITS

The units that cannot be located by field representatives (usually because of inadequate addresses in rural areas) are recorded as unable-to-locate (UTL) units. In a pretest of area segment procedures, conducted in the Charlotte regional office, Harris (1984) reported that 1.6 percent of sample cases (16 out of 1,016) could not be located. The 16 UTL cases came from 12 segments; 4.7 percent of the segments had at least one UTL case. UTL cases also were classified by address information—three post office boxes, seven rural-type address, five multiunit situations and one single-unit in city-type address (that is, street numbers and house number).

The type-A UTL rates experienced in the 1985 AHS-National unit samples from address and area ED's are given in table 4.1 (Schwanz, 1988b). It can be seen that the UTL rates were less than 0.5 percent in address segments and exceeded 2 percent only in area segments in rural areas.

No data on UTL rates for AHS-MS are available but the rates are likely to be similar to the rate inside MSA for the national sample.

Table 4.1. **Unable-to-Locate Rates for the 1985 AHS-National Unit Samples**

(In percent)

Region	Inside MSA				Outside MSA	
	In central city		Not in central city			
	Address	Area	Address	Area	Address	Area
Northeast . .	0.32	0	0.24	0.32	0	2.22
Midwest . . .	0.18	0	0.06	0.12	0	0.41
South	0.22	0.93	0.23	0.52	0.15	0.57
West	0.27	0	0.23	1.69	0.22	2.14

Source: Schwanz (1988b).

ITEM NONRESPONSE

For an interviewed housing unit, information on some items may be missing because of lack of knowledge or refusal by the respondent. If the questionnaire as a whole meets the minimum requirements for a completed interview, missing data for selected items are estimated by imputation. Because AHS is a longitudinal survey, for some missing items we use prior year data, for the rest we use the Census Bureau's traditional sequential ''hot-deck'' procedure (Bailar et al, 1978). The variables used to define imputation matrices vary, depending on the item being imputed. They include race and sex of the reference person, and units in structure. To impute income, AHS uses age, race, and sex of the person, relationship to reference person, and value of property/monthly rent. For each missing value, the procedure substitutes a value reported for a sample unit with similar characteristics. For each item subject to imputation, an indicator variable is added to the AHS data file to show which values have been imputed. The imputed values are, at best, probabilistic in nature, and thus, subject to error, so potential biases from item nonresponse cannot be completely eliminated. Item nonresponse rates vary widely from item to item. Weidman (1988) estimated nonresponse rates for a set of 43 items for the 1985 AHS-National. The results are given in table 4.2. It can be seen that in 1985, 5 out of 43 items had nonresponse rates greater than 10 percent and 12 had rates greater than 3 percent.

Table 4.2. **Nonresponse Rates for Selected Items, 1985 AHS-National**

Items	Nonresponse rate (percent)	Items	Nonresponse rate (percent)
Years on assumed mortgage	25.2	Vandalized	1.1
Amount of mortgage sssumed	23.1	Cost of added/replaced major equipment	.9
Number of toilet breakdowns	13.5	Bars on windows	.9
Amount mortgaged	13.0	Water stoppage	.9
Monthly mortgage payment	11.2	Added/replaced major equipment	.9
Price of home	8.4	Cost of major repairs	.8
Number of water stoppages	8.2	Additions	.7
Maintenance cost in last year	6.1	Square feet of structure	.7
Length of mortgage	5.6	Age of structure	.7
Number of heating equipment breakdowns	4.8	Cost of additions	.6
Reason for insufficient heat	4.5	Undesirable neighborhood	.6
Source of downpayment	3.1	Condition of street	.6
Owned home before	3.0	Public housing	.5
Value of house and property	2.6	Light fixtures	.5
Type of mortgage	2.4	Litter accumulation	.4
Other reason for insufficient heat	2.3	Broken steps	.4
Assumed or new mortgage	2.1	Outside water leaks	.3
Government program mortgage	1.6	Toilet breakdowns	.2
Number of mortgages	1.5	Inside water leaks	.2
Year purchased home	1.4	Value of land	.2
Heating equipment breakdowns	1.3	Insufficient heat	.1
Major repairs over $2000	1.1		

Source: Weidman (1988).

Exhibit 4.1. **Advance Letter to Respondents**

FORM **AHS-66(L)**
(10-13-94)

UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

OFFICE OF THE DIRECTOR

FROM THE DIRECTOR
BUREAU OF THE CENSUS

Your home is among those selected for the American Housing Survey, which the Bureau of the Census is conducting. The Government needs up-to-date facts about the housing in our country to plan Federal housing policy. These data also allow local areas to plan for adequate schools, roads, and other public services. A Census Bureau Field Representative, who will show an official identification card, will call on you within the next few weeks to collect information about your home and your household. Some of the topics covered include the number of rooms, heating and cooling equipment, and the cost of housing.
On the other side of this letter are the answers to questions most frequently asked about this survey.

We would appreciate it if you would look up the following bills in your checkbook or other records before the interview and KEEP THIS LETTER for reference. When the Field Representative calls on you, you can refer to this form. Just skip costs that do not apply to you.

In the past year:

Electric	Gas	for the most recent months of:
$ _____	$ _____	January
$ _____	$ _____	April
$ _____	$ _____	August
$ _____	$ _____	December

$ _____ total cost of fuel oil for the past year

$ _____ total garbage and trash collection costs for the past year

$ _____ total water and sewer costs for the past year

If you own your home:

$ _____ total real estate tax for the past year

$ _____ original mortgage amount

AND current interest rate:

_____ %

Everything you tell us is confidential by law (Title 13, United States Code). We will combine the information you give us with that from many other households so that no one will be able to identify your answers.

Because this is a sample survey, your answers represent not only your home but also hundreds of other homes like yours. For this reason, your cooperation in this voluntary survey will be a distinct service to our country.

This letter is not a questionnaire, but an aid to help you when our Field Representative contacts you. Please DO NOT mail this letter back to us. Keep it for your reference.

Thank you for your cooperation in this essential survey. The Census Bureau appreciates your help.

Sincerely,

Martha Farnsworth Riche

Martha Farnsworth Riche

Exhibit 4.1. **Advance Letter to Respondents**—Con.

What is this survey all about?

The main purpose of the American Housing Survey is to give up-to-date information on the size and composition of the housing inventory. As the Nation grows, so does its demand for housing.

There is a great need for information about the types of homes in which people are now living and about the characteristics of these homes.

Information from the survey is helping to measure the change in our housing resulting from losses and new construction, the structural makeup of the housing, and characteristics of the occupants. It will also help to measure the effect of various tax reform proposals and to generate more mortgage funding into that sector of the economy.

How was I selected for this survey?

Actually, we chose your address, not you personally. The Census Bureau scientifically chose a sample of addresses in metropolitan areas throughout the United States. If you move away, your present address will stay in the survey, and we will interview the family that moves there.

Information about your participation.

Congress requires the U.S. Department of Housing and Urban Development to collect this information under the Housing and Urban-Rural Recovery Act of 1983. The Department of Housing and Urban Development has the authority to collect the survey data under Title 12 of the United States Code and has asked the Census Bureau to conduct the survey. Although there are no penalties for not answering, each missing answer makes the national figures on housing less accurate. The Census Bureau keeps information given us in strict confidence, and requires all Field Representatives to take an oath to uphold and safeguard the confidentiality of all information given them. We will use your answers only for statistical purposes from which no one will be able to identify information about you as an individual.

We expect it will take about 30 minutes for you to provide this information. This time may be somewhat shorter or longer depending on your circumstances. If you have any comments about this survey or have recommendations for reducing its length, send them to the Director, Division of Housing and Demographic Analysis, Office of Policy, Development and Research, Office of Economic Affairs, Department of Housing and Urban Development, Washington, DC 20410; or to the Office of Management and Budget, Paperwork Reduction Project 2528-0016, Washington, DC 20503.

I thought the Census Bureau operated only every 10 years, when it counted people. What is the Bureau of the Census doing now?

Besides the decennial census, which we conduct every 10 years, we collect many different kinds of statistics from other censuses and surveys. We conduct other censuses regularly, including the census of agriculture, the censuses of business and manufactures, and the census of state and local governments. In addition, we collect data on a monthly basis to provide current information on such topics as unemployment rates, retail and wholesale trade, various manufacturing activities, and new housing construction, as well as yearly surveys on business, manufacturing, governments, family income, health, and education. This survey, conducted every 4 years, gives information on homes.

FORM AHS-66(L) (10-13-94)

*U.S.GPO:1994-650-078/00240

Chapter 5.
Measurement Errors

INTRODUCTION

Nonsampling errors, other than coverage and nonresponse errors, that occur during the data collection of the survey are called measurement errors. Some of the potential sources of measurement errors are:

1. Questionnaire design, content, and wording
2. Interview Mode: face-to-face or telephone interview
3. Response errors arising from the respondent due to
    a. Lack of information
    b. Memory problems
    c. Difficulty in understanding questions
    d. Deliberate misrepresentation due to concern over confidentiality or mistrust of the government
4. Interviewer effects. (An interviewer's understanding and participation usually influences responses, especially on questions that have definitional problems.)

QUESTIONNAIRE DESIGN, CONTENT, AND WORDING

Questionnaire structure, content, wording, and sequencing of questions affect responses. Some of the questions tested to develop the questionnaires in field pretest trials are: Do respondents understand questions? Do field representatives (FR's) understand and feel comfortable in asking questions? Does the order of questions influence response? Hayes (1989) reported results of a test of alternative wording for collecting data on electrical breakdowns in the 1988 AHS-MS reinterview. A set of expanded questions resulted in identification of fewer electrical breakdowns than for the regular question. As a result of the changes in the questionnaire in 1985, several items in the 1985 AHS-N and later are not comparable to similar data for 1973 through 1983. Items that changed on the 1985 questionnaire were: units in structure, rooms in unit, plumbing facilities, kitchen, and recent movers. A discussion of each item can be found under the topic of the same name in appendix C of the national report AHS 150/93. The 1995 questionnaire also includes a new sequence intended to improve the collection of information on moderate physical problems, developed as a result of extensive testing (Waite 1993).

INTERVIEW MODE

Telephone interviewing from an FR's home has become an acceptable alternative to personal interviewing in the AHS-National as a result of the telephone experiments conducted in 1981 and 1983. In the 1987 AHS-National, one-third of the sample was assigned to computer-assisted telephone interviewing (CATI) at a centralized facility and the remaining two-thirds continued to be interviewed in the field. CATI experiments continued in the 1989 and 1991 national surveys. Discussions of these telephone experiments and their impact on data follow.

Decentralized Telephone Interviewing Experiments in the AHS-National, 1981 and 1983

Large-scale decentralized telephone interviewing experiments were implemented in conjunction with the 1981 and 1983 enumerations of AHS-National to evaluate the impact of telephone interviewing on data and cost. Parmer, Huang, and Schwanz (1989) analyzed results of these experiments and assessed the impact on data quality and survey costs. Telephone interviewing seemed to have some effect on the data, especially financial characteristics, housing and neighborhood quality characteristics, and income item nonresponse rates. However, this slight effect on data was offset by the cost savings. Overall differences between estimates based on face-to-face interviews and telephone interviews are slightly higher than what could be attributed to random chance. Detailed inspection of the results failed to identify a pattern in the data, however, it was estimated that a 1 percent increase in sample size would make up for the loss in precision due to higher item nonresponse rates. Sample size was not increased, but it was decided that field representatives would conduct telephone interviews whenever possible, beginning with the 1987 AHS-National.

CATI Experiments in the AHS-National, 1987, 1989, and 1991

Computer-assisted telephone interviewing (CATI) techniques may collect data of even higher quality for some data items than do face-to-face or telephone interviews. Using CATI also may help alleviate the effects of the FR staffing retention problems in certain areas by reducing the corresponding field workloads. Therefore, large-scale CATI experiments were implemented in conjunction with the

1987, 1989, and 1991 enumerations of the AHS-National sample to obtain information about the possible effects of CATI on the quality of AHS-National data. Leadbetter et al. (1991) discuss the CATI experiments in 1987 and 1989. Waite (1993) provides the results of the CATI experiments in 1987, 1989, and 1991. In this report, we summarize the results of the CATI experiments as described in Waite (1993).

Design of the experiments. The AHS-National sample is divided equally into six panels. The data of each panel may be used to derive independent estimates of characteristics of interest. Utilizing this feature, two of the six panels were assigned to the CATI treatment and the other four panels were assigned to the non-CATI treatment (personal visit or decentralized telephone interview) in 1987 and 1991. In 1989, four panels were assigned to the CATI treatment and two panels to the non-CATI treatment.

Units assigned to CATI but, for various reasons, not eligible to be interviewed by CATI were screened out and sent to the field for personal visit interviews. The screened units included:

- New construction added since the previous enumeration
- The supplemental sample
- Previous enumeration noninterviews
- Previous enumeration vacant units
- Previous enumeration units temporarily occupied by persons with usual residence elsewhere (URE's)
- Households with 8 or more members
- Multiunit mobile homes
- Units in special places
- Units with address/structure type inconsistencies
- Units interviewed in the previous enumeration that did not have a telephone number where they could be contacted

We considered units not screened out as eligible for CATI and assigned them to the Hagerstown Telephone Center to attempt CATI. CATI eligible units which were not interviewed were recycled to the field for personal visit or decentralized telephone interviews. Therefore, the CATI treatment actually contains units interviewed by all three interview modes.

Preliminary analyses.

How do CATI and non-CATI cross-sectional estimates compare? We compared CATI and non-CATI estimates of household and housing unit characteristics of occupied housing units in our preliminary analysis of the 1987, 1989, and 1991 experiments. We used t-tests to test the hypotheses that estimates from the two treatments were the same. Table 5.1 presents overall results of about 22,000 tests in each year. The overall proportion of significant tests in each year were higher than what would be expected due to chance alone. We expect that 10 percent of the estimates would be different when compared to the α=.10 level of significance, 5 percent at the α=.05 level, and 1 percent at the α=.01 level. Thus, CATI and non-CATI estimates were different indicating that the mode of interview had some impact on data.

Table 5.1. **Proportion of Significant t-tests for CATI Experiments**

(In percent)

Year	Level of significance		
	α=.10	α=.05	α=.01
1987	11.1	6.2	1.9
1989	11.7	6.8	2.3
1991	10.2	5.9	1.7

Source: Waite (1993).

Table 5.2. **Proportion of Significant Differences Between CATI and Non-CATI Estimates**

(In percent)

Subdomains of occupied housing units	Level of significance					
	α=.10		α=.05		α=.01	
	1989	1991	1989	1991	1989	1991
Total occupied	19	15	15	10	6	5
Owner occupied	17	13	11	9	6	5
Renter occupied	14	11	9	6	1	1
New construction	10	10	4	4	1	1
Mobile homes	14	13	8	7	1	2
Severe physical problems	9	9	4	3	1	1
Moderate physical problems	23	11	17	6	8	2
Black	9	10	4	6	1	1
Hispanic	16	7	9	3	3	1
Elderly	12	9	6	6	4	3
Moved in past year	-	10	-	6	-	1
Below poverty level	10	11	5	6	1	2
In MSA's—central cities	11	11	6	7	3	2
In MSA's—suburbs	15	18	10	10	3	3
Outside MSA's	12	6	4	4	1	1
Total urban	20	16	12	10	6	5
Urban—outside MSA's	12	13	7	6	2	2
Total rural	9	8	4	5	1	1
Rural—suburbs	7	10	4	6	1	1
Rural—outside MSA's	10	8	5	4	1	1
Rural—farm	6	5	3	3	1	1
Northeast	9	15	5	9	2	2
Midwest	9	9	5	6	1	1
South	10	10	6	5	1	2
West	13	10	8	6	2	2
Owners with mortgages	-	13	-	11	-	4
Owners with mortgages—specified	-	14	-	11	-	4

- Indicates proportion was not available.
Source: Waite (1993).

The proportions of items with significantly different CATI and non-CATI estimates for various subdomains of occupied housing units in 1989 and 1991 are given in table 5.2. The proportions of significant differences for many items were lower in 1991 than in 1989, as in the case of overall results for all t-tests (table 5.1).

Table 5.3 lists the characteristics with overall significant differences across subcategories and subdomains. These results are from the 1991 experiment. The estimates that displayed differences in 1991 are the same as those with differences in previous experiments, and the direction of the differences for a specific characteristic are generally the same. The direction of the differences are not provided here, but they generally suggest the presence of prestige bias. That is, the CATI treatment had higher estimates of characteristics which suggested that CATI respondents were ''better off'' than did the non-CATI treatment. Non-CATI respondents reported more breakdowns in equipment and more problems in maintaining their housing units.

Table 5.3. **Overall Proportions of Significant Differences for Items Across Subcategories and Subdomains, 1991 CATI Experiment**

(In percent)

Item description	Overall proportion of significant differences α=.10
Cooperatives and condominiums	12.0
Equipment	10.7
Main heating equipment	12.9
Other heating equipment	11.3
Central air conditioning fuel	18.0
Electric fuses and circuit breakers	12.4
Flush toilet breakdowns	26.8
Heating problems	12.2
Water supply storage	14.6
Overall opinion of structure	24.3
Owner or manager on property	10.6
Selected amenities	14.4
Selected physical problems	12.0
Selected deficiencies	10.2
Water leakage during the last 12 months	24.9
Adults and single children under 18 years old	14.7
Household composition by age of householder	10.5
Household moves and formations in the last year	11.4
Location of previous unit	19.0
Persons other than spouse or children	12.0
Persons—previous residence	11.5
Amount of savings and investments	14.4
Food stamps	14.0
Income sources family/primary individual	12.3
Household income	12.0
Rent reductions	16.3
Annual taxes paid per $1,000 value	15.3
Average monthly cost paid for fuel oil	12.0
Monthly costs paid for piped gas	14.0
Monthly costs paid for electricity	15.6
Monthly costs paid selected utilities/fuels	29.0
Monthly housing costs as percent of income	10.5
Other housing costs per month	12.0
Routine maintenance	19.5
Mortgage origination	10.4
Monthly housing costs	10.5
Previous home owned or rented by someone who moved here	10.2

Source: Waite (1993).

Why were CATI and non-CATI treatment data different? Our preliminary analyses were designed to measure the presence or absences of differences between CATI and non-CATI estimates. Following the 1989 preliminary analysis, the 1991 AHS CATI Design Team was formed. One of the objectives of this group was to identify the major reasons for the CATI/non-CATI differences in AHS data. The major findings of this group are listed below:

- Initially, a higher percentage of CATI interviewers did not have survey interviewing experience. In 1989, 40 to 50 percent of the CATI interviewers were new compared to 5 to 10 percent of the field representatives. (None of the 1991 CATI interviewers were new.)
- More non-CATI treatment cases were completed by personal visit. Fifty-six percent of non-CATI treatment cases compared to 45 percent of CATI treatment cases were completed by personal visit in 1989. Fifty-seven percent of non-CATI treatment cases compared to 31 percent of CATI treatment cases were completed by personal visit in 1991. In 1993, 59 percent of non-CATI treatment cases compared to 34 percent of CATI treatment were completed by personal visit.
- Higher item nonresponse rates for some characteristics in CATI. Field representatives probe more often than CATI interviewers. They also have more knowledge about local areas and are better able to recognize when probing is necessary.
- Problems with the probing and interpretation skills of CATI interviewers. This is directly related to the differences in experience, knowledge, and training of the field representatives.
- CATI/non-CATI differences between estimates within the Moderate Physical Problems (MPP) subgroup were probably due to the underreporting of unvented roomheaters by elderly households in the South and in suburbs of metropolitan areas.

To alleviate the effects of the factors listed above, we made several changes to the AHS-National questionnaire and procedures for 1991:

- We improved CATI interviewers training in an attempt to improve field representatives' knowledge of survey concepts and to reduce item nonresponse.
- We added probes to certain items in the CATI questionnaire to automatically appear on the screen if the respondent provided a ''don't know'' response.

- We reconciled the heating equipment and presence of a mortgage item during the CATI interview. We replaced the initial response with the reconciled response whenever the initial response was not correct.
- We planned a reinterview study to determine why there were differences between the number of MPP cases reported by respondents from the two treatments.

The results of most of these changes are discussed in the following section.

Additional analyses.

Nonresponse rates. When respondents do not provide a response for an item, we either allocate a response for the item when we edit the data or publish the total number of nonresponses in ''don't know'' or ''not reported'' categories. The CATI treatment had higher nonresponse for the following items whose nonresponse counts are published:

Lot size
Age of equipment
Overall opinion of structure
Routine maintenance

We compared the allocation (imputation) rates of other items from the 1989 experiment and found the following:

- The CATI treatment had higher allocation rates for items pertaining to financial characteristics (household income, source of income: wages and salaries, monthly cost paid for electricity, water paid separately, trash paid separately, and presence of a mortgage).
- The non-CATI treatment had higher allocation rates for deficiency items (persons per room, lacking complete kitchen facilities, and sewage disposal breakdowns) and general household items (heating equipment and household composition).
- The allocation rates of financial characteristics items were high—averaging about 20 percent. Those for deficiency and general household items were lower (averaging about 3-4 percent).
- Differences in allocation rates between treatments may have produced the differences between the CATI and non-CATI estimates of the following items:

  Steam or hot water heating equipment
  Water paid separately
  Trash paid separately

We added probes to the 1991 CATI questionnaire for the items listed in table 5.4. If a respondent's initial response was ''don't know'' the probes were automatically presented to the field representative in an attempt to get a response to the item. We were not able to add probes for items with higher CATI allocation rates because these results were not available early enough for probes to be incorporated.

Table 5.4. **Item Nonresponse Rates in CATI Experiments**

Item description	Don't know/refused entries			
	1989		1991	
	Number	Percent	Number	Percent
Apartments in building	117	8.76	28	3.08
Age of refrigerator	144	1.64	17	0.28
Age of garbage disposal	92	2.53	21	0.86
Age of oven/burners	131	1.49	25	0.41
Age of dishwasher	64	1.34	11	0.33
Age of washing machine	48	0.64	8	0.15
Age of clothes dryer	35	0.51	7	0.14
Rating of unit as place to live.	74	0.84	50	0.82

Source: Waite (1993).

Table 5.4 illustrates that the addition of probes substantially reduced the proportion of ''don't know'' responses and refusals for the items where probes were added. The reduction in item nonresponse rates can be directly attributed to the computer-controlled interviewing environment of CATI, illustrating that with CATI, we are able to compensate somewhat for interviewer inexperience or poor interviewing skills. This feature is not currently available with decentralized telephone interview or personal visit modes.

Reconciliation study. The AHS survey is longitudinal with the same basic sample being interviewed in consecutive survey years. For items that rarely change, such as whether a housing unit has a basement or not, we would not expect the value of items to change from one enumeration to the next for a particular housing unit. The CATI instrument allows for the storage of prior year data, which can be retrieved for comparison with responses provided during the current interview. The field representative then proceeds with questions based on the results of the prior-to-current year comparisons to determine why there was a change. We took advantage of this feature of CATI in 1987, 1989, and 1991 and reconciled the responses of the following items when the prior and current year responses differed:

Tenure*
Presence of a basement
Number of bedrooms
Number of bathrooms
Type of heating equipment
Type of heating fuel
Amount of rent paid*
Value of home
Whether electricity is included in rent*
Presence of a mortgage*

(The items displayed with an asterisk were not included in the Reconciliation Study in 1989.) The results from the three reconciliation studies were generally consistent, showing that:

- Incorrect responses rather than true changes were the primary reasons for the differences between current and prior year responses for all items except number of bedrooms, value of a home, and presence of a mortgage.
- Unclear definitions for certain items led to erroneous responses.
- Respondents more often reported that the prior, rather than current, year response was wrong.
- Basically, the same items needed reconciliation most frequently each year, signifying respondent reporting difficulties with these items.
- The 1991 CATI and non-CATI treatment responses for most of the reconciled items were significantly different. Number of bedrooms, amount of rent, and presence of a mortgage were treatment-independent.
- There was no difference between the two panels assigned to CATI in 1991. They both have the same distributions for each of the reconciled items. With the exception of the ''value of home'' item, the two panels have the same incorrect response rates.
- The before-reconciliation distributions for the reconciled items were nearly the same as the corresponding after-reconciliation ones.

Although these results add to the mounting evidence of differences between treatments, the studies demonstrate the value of features available to us with CATI but not currently available with the other interview modes—the ability to identify suspicious responses, reconcile them, and make the appropriate changes to the data during the interview, thereby improving data quality.

Gross difference rates. Gross difference rates (GDR's) are indicators of change between consecutive survey years (see Hansen et al., 1964 and Bureau of the Census, 1985b). The reported changes may be due to either response inconsistencies or true status changes. We compared the GDR's of various subdomains and items using 1985 to 1987 longitudinal data. Table 5.5 presents results from the GDR analysis for items like lot size, heating equipment, etc., which were not likely to change from 1985 to 1987. Therefore, GDR's for these items are likely to be close to zero if there was no response inconsistency

Table 5.5. **Results of the Gross Difference Rates (GDR) Analysis of 1985 to 1987 Longitudinal Data, AHS-National**

Subdomain	Item	Results
Total occupied, same households, same owners, all owners, urban, MSA, moderate physical problems (MPP), and below poverty	*Lot size*	Higher CATI GDR's
	Not reported	
	Don't know	
	Main heating equipment	Higher CATI GDR's for:
	Warm air	All except MPP/Below poverty
	Electric heat pump	Urban/MSA only
	Built-in electric units	All except MPP/Below poverty
	Built-in hot air units	All except MPP/Below poverty
	Room heaters without flues	MPP/Below poverty only
	Other portable electric heaters, no heating equipment, or fireplaces without inserts	All except MPP/Below poverty
Same household and same owner	*Main house heating fuel*	Higher CATI GDR's
	Wood	
	Monthly costs as percent of income	Higher CATI GDR's
	Less than 5 percent	
	5-14 percent	
	No cash rent	
	Age of householder	Higher non-CATI GDR's
	45 to 64 years	
	Household income	
	$20,000 to $25,000	Higher non-CATI GDR's for same households
	$25,000 to $40,000	Higher non-CATI GDR's for same households
	$60,000 to $100,000	Higher CATI GDR's for same owners
Same owner	*Routine maintenance in the last year*	
	$75 to $100	Higher non-CATI GDR's
	Not reported	Higher CATI GDR's
	First mortgage payment plan	
	Balloon	Higher non-CATI GDR's
	Other or combination of the above	Higher CATI GDR's

Source: Waite (1993).

Table 5.6. **Results of the Moderate Physical Problems Study**

Reason for the difference	CATI		Non-CATI	
	Number	Percent	Number	Percent
Missed a true MPP	99	37	32	12
Incorrectly counted a MPP	19	7	113	42
Total differences	118	44	145	54

Note: The differences in this table do not total 100 percent. There were eight differences in which both CATI and non-CATI were wrong.
Source: Waite (1993).

between 1985 (all personal interview) and 1987 (CATI and non-CATI). Consequently, higher GDR's reflect more inconsistent responses than lower GDR's. Results show that CATI had higher GDR's for some items and non-CATI for others. Thus, the results tend to indicate that neither CATI nor non-CATI was generally better than the other for producing consistent responses.

Moderate physical problems study. The Moderate Physical Problem (MPP) subdomain had high proportions of significant differences between CATI and non-CATI estimates in both 1987 and 1989. The AHS CATI Design team found that differences between responses to the heating equipment item was the primary reason for these differences. A special MPP study was conducted in 1991 to find out why there were differences within the MPP subdomain. We selected a nonrandom sample of 469 households from cases that were a) interviewed by personal visit in 1991 and b) reported the presence of conditions that would classify the unit as having moderate physical problems. We conducted monitored CATI reinterviews using these cases. The CATI reinterview was an abbreviated form of the AHS-National questionnaire in which all of the questions up to and including the MPP questions were asked. CATI and personal visit interviews were assumed to be two independent responses. If the original and reinterview responses to an item were different, we reconciled the difference to determine why it occurred.

The results of this study indicated a problem in identifying MPP cases in both CATI and non-CATI treatments, but in different directions. Table 5.6 describes how the differences were allocated between CATI and non-CATI. Based on the preliminary results of CATI experiments, we expected the ''CATI/missed a true MPP'' cell would contain the largest proportion of cases in the table below, but were surprised to discover the ''non-CATI/incorrectly counted'' cases were even higher. This means that non-CATI MPP cases were overestimated in addition to CATI MPP cases being underestimated.

Table 5.7. **Responsibility for the Differences in Detecting ''Absolute'' Moderate Physical Problems (MPP's)**

Question	Number of cases	Percent out of total cases[1]	Reconciled cases							Both wrong	Unreconciled cases[3]	
			Percent error of reconciled differences[2]			CATI error		Field error			MPP found in	
			Total reconciled responses	Percent of CATI error	Percent of field error	CATI missed true MPP	CATI incorrectly-counted MPP	Field missed MPP	Field incorrectly-counted MPP		Field not in CATI	CATI not in field
Total	**4246**	**8.5**	**272**	**45.3**	**54.4**	**99**	**19**	**32**	**114**	**8**	**66**	**21**
All toilets broken? (30a)	379	10.6	32	37.5	68.8	8	2	3	17	2	7	1
Fewer than 3 toilets broke for 6 hours? (30b)	2	50.0	1	0.0	100	0	0	0	1	0	0	0
Outside water leaked in? (32a)	380	17.4	52	61.5	40.4	29	2	11	9	1	7	7
Roof/basement leak? (32b=1,2)	87	16.1	9	77.8	33.3	6	0	1	1	1	5	0
Inside water leaked? (32c)	379	19.3	60	45.0	55.0	25	2	3	30	0	12	1
Hot/cold piped water? (33a)	379	1.1	2	0.0	100	0	0	1	1	0	3	0
No running water? (33c)	372	7.3	19	42.1	57.9	7	1	3	8	0	6	2
Connected to a sewer? (35a)	368	1.1	2	100	50.0	1	0	0	0	1	2	0
Sewer broken? (35d)	362	2.5	8	50.0	50.0	4	0	1	3	0	1	0
Unvented heating? (45a=7)	25	64.0	16	56.3	62.5	6	0	0	7	3	0	0
Cracks or holes in walls? (48b)	379	10.0	27	44.4	55.6	10	2	5	10	0	9	2
Holes in floors? (48c)	379	2.4	6	33.3	66.7	0	2	0	4	0	3	0
Peeling paint? (48d)	379	10.3	26	30.8	69.2	3	5	3	15	0	6	7
Rats? (48e)	376	4.8	12	25.0	75.0	0	3	1	8	0	5	1

[1]This column includes reconciled and unreconciled differences.
[2]The differences include when both CATI and non-CATI were wrong.
[3]These cases were unable to be reconciled.

Note: The ''Absolute'' MPP's are the 14 questions that identified whether or not a MPP condition existed.
Source: Waite (1993).

Table 5.7 provides results for various MPP items. The ''water leak'' items (32a, b, c) had the largest number ofCATI versus non-CATI differences of all questions. The ''type of heating equipment'' item accounted for fewer differences, but had the largest rate of difference. Both respondents and field representatives had difficulty understanding the categories of these two items. Some respondents did not know what kind of heating system they had.

Recommendations—Should CATI be used in future enumerations of AHS-National? Although there were still differences between CATI and non-CATI data in 1991, the Census Bureau did not recommend discontinuing CATI for AHS-National. We have identified many positive aspects of CATI. We can continue to use CATI to reconcile questionable results from previous enumerations to improve AHS data quality and as an investigative tool, as we have with the reconciliation and MPP studies. If we interview using CATI in geographic regions with field representative retention problems, we are certain the CATI data we obtain would be no worse than the non-CATI data we would settle for otherwise. Having the ability to monitor and observe inexperienced CATI interviewers while they collect data in these geographic areas is expected to be more desirable than merely accepting data collected in uncontrolled interviews by inexperienced field representatives. Assessment of all these considerations led to a decision to maximize the use of CATI in 1993 and all subsequent survey years.

FIELD REPRESENTATIVE EFFECTS

It is well-known that when a field representative collects data, his/her interaction with respondents and understanding or misunderstanding of questions can have important effects on the results. This is especially the case for questions that are subject to problems with definitions or interpretation. All sample units surveyed by a field representative are subject to correlated field representative effects. This contributes ''correlated response variance'' to the total mean square error in the data.

Such correlated response variance can contribute a substantial portion of the total mean square error for small area population counts and sample estimates in the decennial censuses (Bureau of the Census, 1968). Similar field representative effects on census and survey data have been reported by survey researchers in other countries (Mahalanobis, 1946 and Fellegi, 1974).

An analysis by Tepping and Boland (1972) of interviewer variance in the Current Population Survey gave estimates of 0.50 or greater for the ratio of field representative (correlated response) variance to sampling variance for several items included in the survey. An interviewer variance study carried out in connection with the National Crime Survey in eight cities demonstrated that it can be an important source of error for some variables (Bailey, Moore and Bailar, 1978). The extent to which field representatives influenced crime statistics varied considerably among the cities and among statistics.

There have been no formal interviewer variance studies in connection with AHS. However, the findings from other surveys and from censuses suggest that interviewer variance could be a significant source of error for some items. The contribution of interviewer variance to total error depends on the size of field representative assignments: the larger the assignments, the greater the effect on total error. In both AHS-National and AHS-MS the average field representative assignment is between 30 and 50 households, as compared to about 80 households in the Central Cities Sample for the National Crime Survey. Monthly field representative assignments in the Current Population Survey presently average between 25 and 30 households. In interviewer variance studies cited here, most of the interviews were conducted face-to-face. In AHS-National, where about 50 percent of the interviews are conducted by telephone, the influence of field representatives may be somewhat different.

RESPONSE ERRORS

As mentioned in the section ''Data Collection Instruments,'' chapter 3, page 25, Census Bureau conducts a short second interview called ''reinterview'' within 4 weeks of the first interview, approximately at 20,000 units. By telephone, an experienced field representative tries to talk to the same respondent who talked to the first field representative. Different answers imply that someone made a mistake in at least one of the interviews (or less frequently, that a change has occurred). The rate of discrepancies in response found in the reinterview data over the years are given in table 5.8 for selected items. Information on many more items are given in the AHS Codebook provided to microdata users (HUD and Bureau of the Census, [1973-1993]). It can be seen from table 5.8 that 1 percent of all households changed tenure. In particular, 1 percent of the owners were reclassified as renters, and 2 percent of the renters were reclassified as owners. The two interviews asked about tenure within 4 weeks of each other, so an actual change in tenure would be rare. The differences may be simple misunderstandings. They also may be ambiguous cases (such as a property loaned by a relative, which should be called rental). Note that response errors as indicated by percentage of households changing answers between original interview and reinterview increase with subjective items like street noise, traffic, etc.

Table 5.8. **Response Differences Between Interview and Reinterview in AHS-National**

Item	All units (percent)	Owners (percent)	Renters (percent)	Vacant	Survey year
Different tenure	1	1	2	(NA)	1981
Different occupied/vacant status	3	2	4	4	1981
Different unit visited	.4	-	-	-	1981
Different unit visited	.2	-	-	-	1978
Different household composition	1.0	-	-	-	1981
Different household composition	1.5	-	-	-	1978
Different birthdate	6	-	-	-	1978
Different age	5	-	-	-	1978
Different move date	3	-	-	-	1978

	All[1] (percent)	Yes (percent)	No (percent)	Don't Know	Survey year
Air conditioning	6	7	6	-	1980
To reduce central air use	-	50	-	-	1980
Room unit	1	50	1	-	1980
Awnings	4	50	3	-	1980
Dehumidifier	9	50	5	-	1980
Ceiling fan	5	29	3	-	1980
Attic fan	6	24	5	-	1980
Window fan	4	44	3	-	1980
Portable fan	15	25	12	-	1980
Nothing	23	24	23	-	1980
Added wood/coal stove	3	61	1	-	1980
Added fireplace	1	67	1	-	1980
Added portable electric heater	5	59	3	-	1980
Added unvented kerosene heater	1	86	.3	-	1980
Added other heater	1	69	1	-	1980
Added no heater	10	5	58	-	1980
Have fireplace/stove	6	9	5	-	1980
Fire/stove works	3	2	38	-	1980
All wood bought	14	26	9	-	1980
Had job last week	7	6	7	-	1980
Public transportation besides car	1	55	1	-	1980
Car besides public transportation	7	43	2	-	1980
Same work place daily	5	3	30	-	1980
Garage or carport	5	5	6	-	1978
Piped water in building	40	0	54	-	1977
Had to use extra heat sources	8	44	5	-	1977
Had to use extra heat sources	9	61	5	-	1976
Heating breakdown	6	54	4	-	1977
Heating breakdown	5	40	2	-	1976
Closed unheatable rooms	5	47	3	-	1977
Closed unheatable rooms	4	60	2	-	1976
Interior open cracks/holes	5	49	2	-	1977
Interior open cracks/holes	5	51	3	-	1976
Holes in floors	2	35	1	-	1977
Holes in floors	2	58	1	-	1976
Seen mice or rats	9	40	4	-	1976
Basement	5	5	4	-	1976
Basement leak	15	27	10	38	1976
Electric plug in every room	3	2	49	-	1976
All wiring concealed	3	2	75	-	1976
Attic or roof insulation	28	11	40	55	1976
Thru other bedroom to bath	10	32	5	-	1976
Thru bedroom to other room	6	50	2	-	1976
13 or more shares bedroom with 2 others	19	14	29	-	1976
Blown fuses	10	51	5	100	1976
Garbage collection	7	4	14	100	1976

See symbols and footnotes at end of table.

Table 5.8. **Response Differences Between Interview and Reinterview in AHS-National**—Con.

Item	All[1] (percent)	Yes (percent)	No (percent)	Don't know	Survey year
Mobile home loans	22	17	27	-	1975
Mortgage	1	4	2	-	1975
Water stopped 6 or more hours	13	11	5	75	1975
Roof leaked in last 3 months	5	29	2	42	1974
Roof leaked in last 3 months	5	28	2	51	1973
Main reason for move	15	(NA)	(NA)	-	1973

	All	One	Two	Three	Four or more	Survey year
Number of carpool	17	(NA)	11	37	46	1980
Number of rooms	3	22	30	14	1	1978
Number of bedrooms[2]	6	4	5	6	8	1978
Number of bedrooms[2]	5	6	5	4	7	1977
Heating breakdowns	22	15	40	0	50	1977
Heating breakdowns	26	20	50	25	40	1976

	All	None	One	Two	Three or more	Survey year
Cars owned or used	14	13	10	19	26	1980
Cars owned or used	8	8	5	9	13	1977
Cars owned or used	6	6	4	8	5	1973
Trucks owned or used	9	4	15	37	18	1980
Trucks owned or used	5	3	8	21	-	1977
Rooms without heating ducts	11	5	57	52	29	1977
Rooms without heating ducts	85	6	57	54	34	1976
Blown fuses	17	(NA)	16	30	9	1976

	All	Exclusive use	Shared	No	Survey year
Complete kitchen	1	.3	88	14	1978
Complete kitchen	1	.2	(NA)	26	1977
Complete kitchen	1	.3	89	11	1975
Complete plumbing	1	.2	33	19	1977
Complete plumbing	1	1	46	23	1974

	All[1] (percent)	Excellent (percent)	Good (percent)	Fair (percent)	Poor (percent)	Survey
House rating 2 or more points difference	2	2	.3	4	8	1977
House rating 2 or more points difference	2	2	.4	5	10	1976
House rating 2 or more points difference	1	1	.2	3	10	1975
House rating 2 or more points difference	1	1	.4	2	9	1974
Neighborhood rating 2 or more points difference	2	2	.1	3	39	1977
Neighborhood rating 2 or more points difference	2	2	.4	4	16	1976
Neighborhood rating 2 or more points difference	2	3	0	8	19	1975
Neighborhood rating 2 or more points difference	1	1	.1	2	11	1974
Neighborhood rating 2 or more points difference	1	1	.8	3	1	1973

See symbols and footnotes at end of table.

Table 5.8. **Response Differences Between Interview and Reinterview in AHS-National**—Con.

Item	All (percent)	Have condition (percent)	Do not have condition (percent)	Don't know	All with condition[2] (percent)	No bother (percent)	Little bother (percent)	Much bother (percent)	Want move (percent)	Survey year
Street noise	19	32	14	-	5	5	3	11	10	1977
Heavy traffic	16	27	12	-	-	-	-	-	-	1977
Streets need repair	15	44	8	-	-	-	-	-	-	1977
Snow blocks road	12	48	7	-	-	-	-	-	-	1977
Poor street lighting	17	29	13	-	-	-	-	-	-	1977
Neighborhood crime	12	41	6	-	-	-	-	-	-	1977
Littered streets/lots	13	48	6	-	-	-	-	-	-	1977
Boarded/abandoned buildings	5	31	3	-	-	-	-	-	-	1977
Rundown occupied homes	8	45	5	-	-	-	-	-	-	1977
Nonresidential activities	18	39	14	-	-	-	-	-	-	1977
Odors	8	49	4	-	-	-	-	-	-	1977
Plane noise	13	29	10	-	-	-	-	-	-	1977
Unsatisfactory public transportation	28	31	20	61	-	-	-	-	-	1977
Unsatisfactory schools	14	42	7	50	-	-	-	-	-	1977
Neighborhood shopping	3	43	8	100	-	-	-	-	-	1977
Police protection	85	50	6	68	-	-	-	-	-	1977
Recreation facility	24	43	14	65	-	-	-	-	-	1977
Hospital/clinics	18	48	11	61	-	-	-	-	-	1977

	All renters	Utility paid by household	Included in rent	Not used	Survey year
Different payee for:					
Electricity	2	2	8	0	1981
Gas	13	3	26	20	1981
Other fuels	17	17	47	11	1981
Water	3	10	2	(NA)	1981
Garbage	3	19	1	(NA)	1981

	All owners	Utility paid by household	Not used	Survey year
Electricity	.2	0	40	1977
Gas	1	.5	2	1977

	All[1]	Ducts	Heat pump	Radiators	Built in electric	Floor or wall furnace	Room heaters vented	Room heaters unvented	Fireplace stove or portable	None	Survey year
Main heating	16	11	27	15	13	26	38	21	33	40	1980
Main heating	13	6	53	9	18	26	43	21	28	46	1977
Main heating	7	3	(NA)	4	8	10	19	19	14	18	1975
Main heating	3	4	(NA)	7	8	15	19	14	30	0	1974

	All (percent)	None (percent)	Gas piped (percent)	Gas bottled (percent)	Oil (percent)	Kerosene (percent)	Electricity (percent)	Coal or coke (percent)	Wood (percent)	Solar	Other (percent)	Survey year
Main heating fuel	7	18	5	9	6	27	14	0	17	(NA)	25	1978
Main heating fuel	5	(NA)	3	19	6	50	5	15	16	(NA)	100	1977

See symbols and footnotes at end of table.

Table 5.8. **Response Differences Between Interview and Reinterview in AHS-National**—Con.

Item	All	Wood	Coal	Other	None	Survey year
Fire/stove fuel	9	3	17	25	44	1980

	All	Central	Room units	Survey year
Type of air conditioning	3	2	4	1980

– Not available.
(NA) Not applicable.

[1]"All" means applicable households. For example, piped water was only asked at occupied homes, not vacant.
[2]Different by two or more points.

Source: HUD and Bureau of the Census (1990).

Reinterview data can be used to obtain a statistical measure of discrepancies in responses called "index of inconsistency" (Hansen et al., 1964 and Bureau of the Census, 1985b) defined as:

Index of inconsistent response. This is the ratio of the response variance over the total sampling variance multiplied by 100 so that the index is expressed as a percent. This value can range between zero (indicating perfect consistency between the interview and the reinterview) and 100 (indicating total disagreement between the interview and the reinterview). The following ranges are given as a guideline for interpreting the values:

0 to 20	Low level of inconsistency
20 to 50	Moderate level of inconsistency
50 to 100	High level of inconsistency

L-fold index. This is the same as the above index, but is used for items that have more than two possible answers.

To assess the extent of response errors, the Census Bureau often computes such indices from reinterview data for items that may have problems. A summary of such indices computed from reinterview data from 1973 through 1985 has been compiled by Chakrabarty (1992a). Table 5.9 provides L-fold indexes for selected AHS items. It can been seen that opinion questions like adequacy or inadequacy of recreation facilities, and items that are not easy to remember like the number of electrical blowouts in last 90 days, have a high level of inconsistency.

Besides regular reinterviews, the Census Bureau conducts periodic studies to determine the extent of response problems. In the 1987 AHS-National Survey, the answers to selected questions provided by households interviewed by CATI were compared to the answers provided by the same respondents in 1985. If the answers were different the field representative asked the respondent to explain the discrepancies. This was done immediately after the completion of the 1987 interview while the respondent was still on the telephone. The results of this study using a sample of 6,268 households reported earlier in HUD and Bureau of the Census (1990) are presented here in table 5.10.

Table 5.9. **Index of Inconsistent Responses Between Interview and Reinterview for Selected Items, AHS-National**

Item	L-fold index	Survey year
Is public transportation? Adequate, inadequate, ...enough to move, don't know (DK)	49.8	1977
Are the schools? Adequate, inadequate, ...enough to move, DK	49.8	1977
Is the shopping (drug stores and grocery stores)? Adequate, inadequate, ...enough to move, DK	53.5	1977
Are the outdoor recreation facilities (parks, playgrounds, etc.)? Adequate, inadequate, ...enough to move, DK	54.6	1977
How many of the neighborhood services are inadequate? One or more, none, DK	55.1	1976
How many are so inadequate that the respondent would like to move? One or more, none, blanks	47.3	1976
Does your house/apartment have garbage collection service (public or private)? Yes, no, DK	17.6	1976
How do you dispose of garbage? Incinerator, trash chute or compactor, put out to pickup, other	66.3	1976
Have any electric fuses or breaker switches blown in your house/apartment in the last 90 days? Yes, no, DK	58.0	1981
How many times did this happen? Once, twice, three or more	50.0	1981
What type of heating equipment does your house/apartment have? Central air, heat pump, steam system, etc.	25.1	1981
What heating fuel do you use? Gas from pipes, gas from tank, fuel oil, kerosene, electricity, coal or coke, wood, etc.	10.8	1978
How many rooms are in this house/apartment (don't count bathrooms, porches, halls, or half-rooms)? One, two, three, four or more	13.8	1978
How many bedrooms are in the house/apartment? One, two, three, four or more	7.5	1985
How much do you think this property; that is, house and lot, would sell for on today's market? Less than $5,000, $5,000-7,499, $7,500-9,999, ...$200,000 or more	31.9	1981
In regard to the mortgage, what is the amount of the required payments to the lender? Less than $100, $100-149, $150-199, ...$1,000 or more	19.0	1981
What is the yearly cost of your real estate taxes? Less than $100, $100-199, $200-299, ...$1,000 or more	39.8	1981
What is the yearly cost of your fire and hazard insurance? Less than $40, $40-59, $60-79, ...$180 or more.	47.3	1981
In view of all things discussed, how would you rate this street as a place to live? Excellent, good, fair, poor	47.9	1977
In view of all things discussed how would you rate this house/building as a place to live? Excellent, good, fair, poor	45.6	1977

DK Do not know.
Source: Chakrabarty (1992a).

Table 5.10. **Reasons for Discrepancies Found Between 1985 and 1987 Out of 6,268 Households Examined, AHS-National**

TENURE	Reason
Purchased since 1985	21
Sold, now renting	4
Began charging rent since 1985	1
Stopped charging rent since 1985	2
1985 answer wrong	42
1987 answer wrong	41
Other	38
Total	**149**

BASEMENT	Reason
Built under house	3
Old basement filled in	1
House is split-level, don't know what to call it	17
Have a partial basement, don't know what to call it	18
Walkout basement, don't know what to call it	0
Shallow basement, don't know what to call it	2
1985 answer wrong	305
1987 answer wrong	349
Other	60
Total	**755**

BEDROOM	Reason
Another room converted	144
Addition added	34
Bedroom now used for something else	219
Part of house/apartment merged	4
Attic or basement finished	19
1985 answer wrong	127
1987 answer wrong	164
Other	61
Total	**772**

BATHROOM	First reason	Second reason
Half converted to full	15	0
Added in addition	52	0
Space converted	7	0
Some/all fixtures removed	5	0
Destroyed in merger	0	0
1985 answer included half bathrooms	6	1
1987 answer included half bathrooms	6	0
1985 answer wrong	253	4
1987 answer wrong	152	1
Other	29	2
Refused	1	-
Total	**526**	8

FUEL	Reason
Fuel used less often in 1985, now more	152
New/converted equipment used other fuel	87
1985 answer wrong	133
1987 answer wrong	155
Other	83
Refused	4
Total	**614**

HEATING EQUIPMENT	First reason	Second reason
Old equipment replaces	80	0
Types used less 1985, now more	150	3
Installed since 1985	36	1
1985 answer wrong	359	2
1987 answer wrong	480	2
Other	80	5
Refused	11	-
Total	**1,196**	**13**

RENT	Paid monthly first reason	Paid monthly second reason	Paid yearly first reason	Paid yearly second reason
Major alterations/ improvements	6	0	1	0
Conversion or merger changed size of unit	0	0	0	0
Disaster/partial demolition changed	0	0	0	0
No longer rent controlled	1	0	0	0
Now rent controlled	1	0	0	0
No longer subsidized	1	0	0	0
Now subsidized	6	0	0	0
Owner raised/lowered rent	76	0	5	0
1985 answer wrong	12	5	4	1
1987 answer wrong	10	0	4	1
Other	33	5	3	2
Refused	1	1	1	0
Total	**147**	**11**	**18**	**4**

VALUE	First reason	Second reason
Major alterations/improvements	89	13
Disaster/demolition	0	1
Sold/purchased land	3	0
Area more developed	68	23
Area had major disaster	3	1
Changes in the economy	253	54
Rezoning	4	1
1985 answer wrong	296	7
1987 answer wrong	77	4
Other	190	25
Refused	9	1
Total	**991**	**130**

Source: HUD and Bureau of the Census (1990).

1985 AHS-MS REINTERVIEW

In 1985, reinterview measured response variance of selected questions not previously evaluated. Survey items reviewed generally fall into three categories: (1) mobility, (2) major repairs, and (3) mortgage. The major results of the analysis of reinterview data are given below from Waite (1990b).

Reasons Moved (Question 52)

For the first part of this question, the respondent was to indicate all categories that apply. The question had 15 categories. For analysis we created a mention/did not mention table for each category. This question was tallied when at least 1 of the 15 categories was marked in both interviews.

Of the 15 categories, 6 showed high response variance, 6 showed moderate response variance, none showed low response variance, and 3 did not meet the minimum requirements necessary to compute reliable estimates of the index. The categories and their indexes are listed in table 5.11.

The second part of this question asked ‘‘What is the MAIN reason you moved?’’ of all the 15 categories plus an additional category of ‘‘all reasons of equal importance.’’ The question showed high response variance. Results are given in table 5.11.

These results suggest that the question needs improvement to produce more reliable data.

Table 5.11. **Response Variance: Reasons Moved (Questions 52a and 52b) 1985 AHS-MS**

Reasons moved (question 52a)	Index of inconsistency
1 Private company or person wanted to use	*
2 Forced to leave by government	*
3 Disaster loss (fire, flood, etc.)	*
4 New job or job transfer	30
5 Be closer to school/work	41
6 Other, financial/employment related	80
7 To establish own household	48
8 Needed larger house or apartment	33
9 Married, widowed, divorced, or separated	35
10 Other, family/personal related	68
11 Wanted better quality house	69
12 Change from owner to renter OR renter to owner	44
13 Wanted lower rent or less expensive house to maintain.	55
14 Other, housing related	79
15 Other	73
Main reason moved (question 52b)	51 (L-fold)

* Not enough sample cases to compute a reliable estimate of the index.

Source: Waite (1990b).

Major Repairs

The AHS asked a set of three questions about nine different major repairs, improvements or alterations made to the house/apartment in the last 2 years.

1. Was the (repair) done?
2. Did someone in the household do most of the work?
3. How much did the job cost (not counting household members’ time)?

For the first question all repairs except one had moderate response variance. The exception was the catch-all category, ‘‘Any (other) repairs over $500,’’ which had high response variance.

For the second question, one repair had high response variance, one had moderate response variance, three had low response variance, and four did not meet the minimum requirements necessary to compute reliable estimates of the index.

For the third question, the only repair that met the minimum requirements to compute a reliable estimate of the index had low response variance. We used three cost categories for this analysis: no cost, less than $500, and greater or equal than $500. Table 5.12 shows the categories and the indexes.

Table 5.12. **Response Variance: Major Repairs (Question 73) 1985 AHS-MS**

Type of repair	Index of inconsistency		
	Repair done (a)	Someone in household do work (b)	L-fold index—Job cost (c)
1 All or part of roof replaced in last 2 years	[1]35	25	*
2 Any additions built	46	*	*
3 Kitchen remodeled or added	32	9	*
4 Bathrooms remodeled or added	35	*	*
5 Siding replaced or added in last 2 years	42	*	*
6 New storm doors or storm windows bought and installed	33	19	15
7 Major equipment, such as furnace or central air replaced or added	44	*	*
8 Insulation added	32	16	*
9 Other major repairs over $500 each	57	51	*

* Not enough sample cases to compute a reliable estimate of the index.

[1]This is an L-fold index—this question had three response categories: yes all, yes part, no.

Source: Waite (1990b).

Mortgage (Question 96)

The mortgage question group contained a series of 29 questions asking if respondents had a first mortgage and repeated these questions for second mortgages. None of the second mortgage questions met the minimum requirements necessary to compute reliable estimates of the index. One question of the first mortgage group had high response variance, 6 had moderate response variance, 7 had low response variance, and 15 did not meet the minimum requirements necessary to compute reliable estimates of the index. Table 5.13 shows the questions and the indexes.

Mobility Supplement (Questions 177-183)

The mobility supplement asked questions on people's moving patterns. Three questions had high response variance, three had moderate response, and one did not meet the minimum requirements necessary to compute a reliable estimate of index. Table 5.14 shows the questions and indexes.

Table 5.13. **Response Variance: First Mortgage (Question 96) 1985 AHS-MS**

	Mortgage question	Index of inconsistency
a	Current mortgage same year as bought home	39
b	New or assume someone else's	15 (L-fold)
c	Amount left to pay off when you assumed it	* (L-fold)
d	How many years remained on mortgage then	* (L-fold)
e	What year get mortgage	* (L-fold)
f	When first obtained THIS mortgage, how many years was it for	22 (L-fold)
g	At current payments, how long to pay off loan	* (L-fold)
h	How much was borrowed	14 (L-fold)
i1	Mortgage cover other homes or apartments	*
i2	Mortgage cover farm land	*
i3	Mortgage cover a business on this property	*
j	How much applies just to your home	* (L-fold)
k	Current interest rate on mortgage	12 (L-fold)
l	Current monthly payment	10 (L-fold)
m1	Payment include property taxes	18
m2	Payment include homeowner's insurance	18
m3	Payment include anything else	48
m4	How much were other charges last year	* (L-fold)
n	Type of mortgage	21 (L-fold)
o	Borrow money from a bank or other organization OR borrow from an individual	11
p	Borrow from the former owner of home	*
q	Payments the same during whole length of the mortgage	52
r1	Change in taxes or insurance, or due to decline in principal balance	37
r2	Change based on interest rates	26
r3	Rise at fixed schedule during part of loan	*
r4	Rise at fixed schedule during whole length of loan	*
r5	Last payment biggest	*
r7	Other change	*
r8	Of total amount borrowed, what percentage will have to be payed off in last payment	* (L-fold)

*Not enough sample cases to compute a reliable estimate of the index.
Source: Waite (1990b).

Table 5.14. **Response Variance: Mobility Supplement (Questions 177-183), 1985 AHS-MS**

	Mobility question	Index of inconsistency
177a	At age 16, live in this area or a different place	50
177c	Which best describes place above AT THAT TIME	51 (L-fold)
178	Five years from now, PREFER to be living in this house/apartment or someplace else	24
179	Five years from now, how LIKELY to be living in this unit	60 (L-fold)
180	Five years from now, prefer to be living in another home in this area or outside this area	38
181	Which best describes the area would prefer to live in 5 years from now	* (L-fold)
183	Within next 5 years, how LIKELY to move to place prefer to live	62 (L-fold)

*Not enough sample cases to compute a reliable estimate of the index.
Source: Waite (1990b).

Conclusion

To produce more reliable data, the questionnaire needs some modification. Specific recommendations for some of the high response variance questions include:

- Explore ways to reword question 52b (reasons moved).
- Provide clearer definitions of city size; that is, give population ranges and provide flashcard in question 181 of the Mobility Supplement.
- Combine the categories ''Very Likely'' and ''Likely'' in questions 179 and 183 of the Mobility Supplement.

The lot size question needs more responses for evaluation. Forty-five percent of the usable responses were ''Don't Know'' in either the reinterview or the original interview. Of the number responses given, 51.5 percent agreed in both interviews while 79 percent agreed within +/– 20 percent of each other.

RESPONSE ERROR IN YEAR BUILT DATA

Stating the year in which the structure was built has always been a problem for respondents in the AHS and other surveys; for example, CPS, and in the census as well, particularly when he/she is not the first owner of the housing unit or when the respondent is renting rather than buying. This problem has been reported in many studies. In this report, we provide a summary of results of the studies conducted by the Bureau of the Census.

A content reinterview for the 1980 census showed that the year built data have considerable response variance and bias (overreporting or underreporting). The multiunit structure data displayed higher response variability and bias than the single unit data. Also, the response variability in the year built data in the 1980 census was at about the

same level as in the 1970 census (see, Bureau of the Census, 1986). Similar reinterview data from the AHS (National or MS) are not available.

The ''year built'' item was one of two items selected for a record check in the ''Tampa AHS Census Match Study'' (Tippett, 1988). Since suitable administrative records were located at the County Tax Assessor's office, a check was performed to see if the item was effectively measuring the characteristic of interest. The results of the record check are shown in table 5.15 for responses that were given in the census.

Table 5.16 shows the responses given in the AHS that were compared to the assessor's file.

Table 5.15. **Frequency Distribution of Census Households by Year Built From the Tampa AHS Census Match Study, 1985**

Year built	Source			
	Census		Assessor's files	
	Number of units	Percent	Number of units	Percent
1980 or later	5	2.3	7	3.2
1970-79	32	14.8	31	14.4
1960-69	59	27.3	47	21.8
1950-59	62	28.7	69	31.9
1940-49	36	16.7	36	16.7
1939 or earlier	22	10.2	26	12.0
Total answering item	**216**	**[1]100.0**	**216**	**100.0**

[1]Fourteen (6.1 percent) out of the 230 respondents whom had their records checked against the assessor's file left the question blank.

Source: Tippett (1988).

Table 5.16. **Frequency Distribution of AHS Households by Year Built From the Tampa AHS Census Match Study, 1985**

Year built	Source			
	AHS		Assessor's files	
	Number of units	Percent	Number of units	Percent
1980 or later	4	2.2	7	3.9
1970-79	30	16.9	35	19.7
1960-69	42	23.6	42	23.6
1950-59	49	27.5	47	26.4
1940-49	27	15.2	24	13.5
1939 or earlier	26	14.6	23	12.9
Total answering item	**178**	**[1]100.0**	**178**	**100.0**

[1]Fourteen (7.3 percent) out of the 192 AHS respondents whom had their records checked against the assessor's file left the question blank or answered ''Do not know.''

Source: Tippett (1988).

Table 5.17 shows the percentage of the respondents from the AHS and census whose responses agreed with the tax assessor's files. It also shows a breakdown for owners versus renters. Note that the percentages are calculated taking into consideration only those households that responded to the question, so a household's response can only be in disagreement with the assessor's file if a response is given to the question and that ''year built'' was not the same as the assessor's ''year built.''

It can be seen that the overall agreement of responses with the assessor's file is about the same for both census and AHS respondents. Owners naturally had better information on when the unit was built compared to renters in the census. The high (14.5 percent) nonresponse rate for renters (see table 5.18) in this study for AHS might have biased the result. In any case, the differences between owners and renters based on a small AHS sample were not statistically significant.

A separate study was done in Tampa to see if data on multiunit structures (rental units) from knowledgeable sources such as a landlord, property manager, realtor, etc. was more reliable. Eight (7.6 percent) of the 113 informed structure respondents whose responses were checked against the assessor's file did not answer or answered ''Don't know.'' Out of the 105 cases where the respondent actually answered the question, 97 (that is, 92.4 percent) cases agreed with the assessor's file.

Tippet concluded, ''There is not so much a reluctance to answer this item, as that respondents simply do not have the information to answer accurately and are giving their ''best guess''. . . . An informed structure respondent was a much better source compared to renters.'' It was recommended that further structure respondent testing be done.

The problem of the ''year built'' item also was addressed by Young (1982). Table 5.19 provides year built for all year-round units (in thousands) for the 1980 census and AHS-National.

Table 5.17. **Percentage in Agreement With Assessor's File**

Source	All	Owners	Renters
Census	80.6	83.7	73.0
AHS	83.1	80.7	88.1

Source: Tippett (1988).

Table 5.18. **Nonresponse Rate (Percent) for the "Year Built" Question**

Source	Overall	Owners	Renters
Census	6.1	5.0	8.7
AHS	7.3	3.3	14.5

Source: Tippett (1988).

Table 5.19. **Comparison of Year Built Data for All Housing Units in the 1980 Census and AHS-National**

(In thousands)

Year built	Housing units	
	1980 census	1980 AHS
1970-80	22,434	19,735
1979-80	2,926	3,433
1975-78	8,381	7,071
1970-74	11,126	9,231
1960-69	16,861	17,624
1950-59	14,995	14,043
1940-49	9,813	7,945
1939 or earlier	22,667	26,677
Total	**86,769**	**86,024**

Note: Census estimates include 510,000 vacant year-round mobile homes, while the AHS estimates do not.

Source: Young (1982).

It can be seen that several discrepancies exist between AHS and census estimates. A difference of 2.7 million units for the 1970-80 cohort is most striking. Young states that, ''there are several possible reasons for the 1970-80 cohort difference of 2.7 million units:

- A potential response error problem in the census. We know from past experience (1970 census evaluation program) that this is a problem.
- An excessive number of erroneous inclusions in the census; for example, duplicates, erroneous enumerations, etc. that were built during the period 1970-80.
- Serious undercoverage problems in the AHS of units built during the period 1970-80.''

As stated at the beginning, the ''year built'' question has always been a problem, not just for the AHS, but other surveys and the census as well. The problem is not so much with the surveys themselves as with the content of the question. Some respondents seem to have a lack of genuine knowledge about the structure in which they live or a poor understanding of what the question is asking. The respondent fails to realize that the question is being asked about the structure and not the unit in which he/she resides, and the year the structure was built is not determined by a conversion, rehabilitation, redecorating, or additions to the unit. Poor coverage of certain types of structures also contributes to unreliable counts (although this may be improving).

Perhaps a more clearly asked question, better training of field representatives for this item, and a more informed respondent (for example, a landlord, property manager, realtor, etc.) would help to increase the reliability of this question.

Finally, note that owners are a more reliable source of information than renters in that renters had a higher rate of ''don't know'' responses.

PROBLEMS WITH THE NUMBER OF UNITS IN STRUCTURE QUESTION

The number of units in a structure is a basic housing characteristic. A respondent is asked how many units there are in the structure in which his/her unit resides. A distinction is made between a housing unit; for example, an apartment, townhouse, condominium, and the structure in which the unit is contained. The structure or building may consist of one or many units. Furthermore, single unit structures are classified as either detached or attached to other structures. This question seems to give respondents a conceptual problem, especially in classifying townhouses, duplexes, and small attached units and in making a distinction between a housing unit and a structure. Taeuber, et al (1983) compared 1980 census estimates of the totals of the ''units in structure'' categories with AHS-National estimates. These results are shown in table 5.20.

The differences, except the totals, are greater than those expected from sampling error. Since the census was taken as of April 1, 1980, and the AHS date was October 1980, the total estimate of housing units is expected to be 800,000 to 1,000,000 units higher in the AHS than in the census due to new construction. This is not the case however; the increase was only 335,000 units. It can seen that the most notable difference existed in the ''5 or more units'' category.

Young (1982), who also examines the problem, states that the possible reasons for this discrepancy are:

- ''Census misclassification error. There has been some concern that census respondents might have incorrectly identified certain types of single (or 2-to-4-unit structures) as 5 or more structures; for example, attached townhouses or garden apartments.
- Serious undercoverage problems may exist in our current surveys for picking up new large multiunit structures.''

Table 5.20. **Units in Structure of All Housing Units From the 1980 Census and AHS-National**

(In thousands)

Units in structure	Housing units	
	1980 census	1980 AHS
1	57,183	58,255
2-4	9,682	10,816
5 or more	15,478	13,183
Mobile home	[1]4,416	[1]4,840
Total	86,759	[1]87,094

[1]Data adjusted to include vacant mobile homes for comparability with the 1980 census.

Source: Taeuber, et al. (1983).

Table 5.21. **Units in Structure of Occupied Housing by Tenure From the 1980 Census and AHS-National**

(In thousands)

Units in structure	Housing units			
	Renters		Owners	
	Census	AHS	Census	AHS
1	8,731	8,558	45,130	46,330
2 to 4	6,668	7,468	2,167	2,247
5 or more	12,415	10,801	1,404	897
Mobile home	779	728	3,095	3,041
Total	**28,593**	**27,556**	**51,796**	**52,516**

Source: Young (1982).

Young also provides units in structure data cross-classified by owners and renters. Table 5.21 compares the estimates for renters and owners.

It can be seen that estimates for 1 unit and 2 to 4 units are remarkably close considering the time differential between the census and AHS. Most of the discrepancy in estimates is due to the ''5 or more units'' category. The AHS seems to have coverage problems for structures with 5 or more units and for within structure conversions.

The ''units in structure'' problem also was studied by Tippett (1988). Table 5.22 provides a comparison of census and AHS responses in more detail than in previous studies (that is, there are more units in structure categories).

The responses in the above table are displayed in table 5.23 to show disagreement between AHS and census responses, and need for reconciliation. The results further demonstrate the problem of classification in moderate-to-large sized buildings. It shows that of the 399 units for which a response was recorded for both the census and the AHS, 304 (on the diagonal) agreed and 95 had conflicting responses. Forty-nine of these ninety-five responses were reconciled. Some of the difficulties which respondents encountered included the following: ''buildings were connected at the roof line or via connecting pathways outside the upper floors, so that it was difficult to determine exactly where one building ended and the next one began'' and ''respondents gave counts for their apartment complex rather than for their particular apartment building.'' These complications and others, along with the index of inconsistency of 35.15, indicate that there may be a problem with the reliability of the data collected with this item.

Table 5.22. **Units in Structure for All Housing Units in the Tampa AHS Census Match Study, 1985**

Units	Test census		AHS	
	Number	Percent	Number	Percent
1 detached	220	51.4	222	51.9
1 attached	16	3.7	8	1.9
2	23	5.4	25	5.9
3 to 4	24	5.6	24	5.6
5 to 9	32	7.5	37	8.6
10 to 19	25	5.9	45	10.5
20 to 49	18	4.2	25	5.8
50 or more	43	10.0	17	4.0
Mobile home	11	2.6	9	2.1
Reported (subtotal)	412	96.3	412	96.3
Not reported	16	3.7	16	3.7
Total	428	100.0	428	100.0

Source: Tippett (1988).

Table 5.23. **Units in Structure by the Test Census and AHS Responses, Tampa Study 1985**

AHS	Census									
	1 detached	1 attached	2	3 to 4	5 to 9	10 to 19	20 to 49	50 or more	Mobile home	Total
1 detached	205	3	3	1	0	1	0	0	4	217
1 attached	3	4	1	0	0	0	0	0	0	8
2	4	3	16	0	0	0	0	0	0	23
3 to 4	1	1	0	16	0	3	1	2	0	24
5 to 9	1	1	2	4	19	6	0	3	0	36
10 to 19	0	1	1	1	9	11	6	12	0	41
20 to 49	0	1	0	0	3	3	10	8	0	25
50 or more	0	0	0	0	0	1	0	16	0	17
Mobile home	0	1	0	0	0	0	0	0	7	8
Total	**214**	**15**	**23**	**22**	**31**	**25**	**17**	**41**	**11**	**399**

Note: The elements along the diagonal of this table indicate the number of units whose responses agreed in both the test census and the AHS. The index of inconsistency for the table, after reconciling the results, was 35.15, which is considered to be moderate.

Source: Tippett (1988).

Finally, we will consider a study described by Abernathy (1987) for the 1987 AHS-MS. The responses from Wave I of the Regional Office Preedit were compared to the responses from the last enumeration period for AHS. This is part of the continuing quality control program which checks for and corrects inconsistencies. When the ''units in structure'' response is found to be inconsistent with the previous answer, the response is flagged. Table 5.24 provides a distribution of 119 rejected records by type of inconsistency.

The two main types of inconsistencies are as follows: ''units that were classified as one attached 1 year and in a multiunit structure the other year; and units that were classified as in multiunit structures both years, but the number of units in the structure between survey years was inconsistent.''

Also, part of the quality control process was not only to detect the types of inconsistencies with the previous year, but also to check the corrected responses with the previous year. In other words, once the correction cycle is run on the data that are flagged as ''units in structure inconsistent,'' the responses are again checked with the entries from the previous enumeration period. At this point it has been determined that the majority of the corrected entries are consistent with the prior year's entries. Abernathy concludes, ''it appears that the preedit research is doing its job in reducing the classification problems that exist with the current year's data.''

There have been several improvements to both the census and AHS approaches. Since the Tampa Match Study, ''the census item has been clarified to refer to the building rather than the address, and in the AHS, examples in field representative training are more explicit in pointing out the problems in classification.'' Also, it was suggested that ''the confusion that exists in some instances also highlights the importance of supplementing address lists for multiunit structures with an on-the-ground reconnaissance by an enumerator during field operations, such as prelist and precanvass in improving unit identification in multiunit situations.''

However, there are still problems with the ''units in structure'' item in the AHS. First, there is an indication of a coverage problem for moderate to large structures, more specifically in the ''five or more'' units category. This problem also may be caused by poor coverage of structures that have had conversions done to them. Secondly, the question still seems to be conceptually difficult for the respondents. They do not fully realize the definitional differences between their units and the structures in which those units are located. This becomes more difficult for them when the structure is attached to another structure.

Table 5.24. **Inconsistencies in the Units in Structure Data Compared to the Previous Response, 1987 AHS-MS**

Current year response	Prior year response	Total
1 detached	Multiunit	3
1 attached	Multiunit	32
Multiunit	1 detached	8
Multiunit	1 attached	15
Multiunit: more units	Multiunits: fewer units	19
Multiunits: fewer units	Multiunits: more units	23
Multiunits: number left blank	Multiunit	19
Total		**119**

Source: Abernathy (1987).

PROBLEMS WITH THE TENURE QUESTION

Tenure is important as a basic housing characteristic. The tenure question asks the respondent if he/she owns the unit, rents for cash, or occupies without payment of cash rent. The tenure question presents few conceptual problems for respondents, but the owner occupancy rates are persistently higher in surveys than in the census. This fact is documented by Taeuber, Thompson, and Young (1983) in the report ''1980 Census Data: The Quality of Data and Some Anomalies,'' where they compared the owner occupancy rate of the census with that from the Current Population Survey/Housing Vacancy Survey.

Table 5.25 provides a comparison of census and AHS responses for tenure from the ''Tampa AHS Census Match Study'' by Tippett (1988).

It can be seen that the AHS does have a slightly higher occupancy rate for owners than does the test census. These figures can be further broken down to show the responses that did not agree between the two sources, and needed to be reconciled. This is done in table 5.26.

It can be seen that of the 324 respondents who replied to both the test census and the AHS, 304 agreed and 20 gave conflicting responses. Thirteen of those 20 responses were reconciled. During the reconciliation, reasons for the discrepancies were discovered and listed in the report as follows: ''for two cases, a change of tenure had occurred, so both were correctly enumerated; others resulted from mismarking of the item, different respondents, or a temporary interruption in the rent.'' These incidental discrepancies are not indicative of any problem that is inherent in the

Table 5.25. **Tenure Responses for All Occupied Units in the Tampa AHS Census Match Study, 1985**

Characteristics	Test census		AHS	
	Number	Percent	Number	Percent
Owned	158	42.2	168	44.9
Rented for cash	200	53.3	197	52.7
Occupied without payment of cash rent	2	0.5	5	1.3
Reported (subtotal)	360	96.0	370	98.9
Not reported	15	4.0	4	1.1
Total occupied units[1]	375	100.0	374	100.0

[1]The totals do not match because one of the units which was occupied during the test census was not occupied during the AHS.

Source: Tippett (1988).

Table 5.26. **Tenure Responses for Occupied Housing Units by the Test Census and AHS, Tampa (1985)**

Census	AHS			
	Owner	Rented for cash	Occupied without payment of cash rent	Total
Owner	139	8	0	147
Rented for cash	8	164	0	172
Occupied without payment of cash rent	1	3	1	5
Total	**148**	**175**	**1**	**324**

Source: Tippett (1988).

tenure question, and they do not help to explain the problem of the differences in the owner occupancy rates between the census and the AHS.

As an additional note, once the results have been reconciled the tenure item has an L-fold index of inconsistency in the low range, 11.08. This indicates that the respondents are answering the tenure question reasonably well.

VERIFICATION OF REPORTING OF COOPERATIVES AND CONDOMINIUMS

To evaluate the accuracy of the classification of housing units as cooperatives and condominiums in the AHS-National, part of the reinterview program for 1979 focused on verifying responses to the AHS questions on cooperative and condominium status.

Completed AHS questionnaires were screened in the regional offices to identify questionnaires where the response indicated that the unit was a cooperative or condominium.

A followup interview was conducted by the field staff to ascertain whether the original response was correct.

Followup was generally done by telephone and attempts were made to interview a knowledgeable respondent such as a building manager, sales agent and the like. If no such respondent could be located, the respondent to the original AHS interview was reinterviewed. In addition to verifying cooperative or condominium status, some additional questions on conversion and date of conversion were asked during the followup interview.

Verification showed that out of 935 units originally classified as condominiums, 24 (2.6 percent) were not condominiums, and out of 159 units classified as cooperatives, 50 (31.4 percent) were not cooperatives (Buckles, 1981). Reporting of cooperative status was much less accurate than the reporting of condominium status. A cooperative probe was developed from this verification to clarify the cooperative definition for the respondents. The probe was added to the questionnaire in 1980 and has been used ever since in the survey.

Another verification of cooperatives and condominiums, similar to the 1979 verification was conducted in 1983. Hartnett (1985) provided results of this study.

Condominiums

The verification followup showed that of the 1,634 units originally reported as condominiums, 62 units had a status that changed to not condominium (see table 5.27). The original count exceeded the verification followup by 3.8 percent; this was a larger proportion than the 2.6 percent discovered in a similar study done in 1979. However, because the instructions were misinterpreted in the 1979 study, many renter-occupied condominiums were not followed-up. The explanations for misclassification are listed in table 5.28.

Table 5.27. **Verification Results for Housing Units Originally Classified as Condominium: 1983 AHS-National**

Item	Number	Percent
Total (occupied and vacant)	**1,634**	**100**
Verified as condominium	1,371	83.9
Verified as not condominium	62	3.8
Followup not completed	201	12.3
Owner-occupied	**816**	**100**
Verified as condominium	678	83.1
Verified as not condominium	22	2.7
Followup not completed	116	14.2
Renter occupied	**397**	**100**
Verified as condominium	329	82.9
Verified as not condominium	28	7.0
Followup not completed	40	10.1
Vacant for sale	**101**	**100**
Verified as condominium	81	80.2
Verified as not condominium	0	0
Followup not completed	20	19.8
Vacant for rent	**320**	**100**
Verified as condominium	283	88.4
Verified as not condominium	12	3.8
Followup not completed	25	7.8

Source: Hartnett (1985).

Table 5.28. **Reasons Condominiums Were Misclassified 1983, AHS-National**

Reason	Number	Percent
Total	**62**	**100**
Unit was cooperative	10	16.1
Townhouse, not a condominium	11	17.7
"Renter not knowing condominium definition"	4	6.5
Homeowner association	6	9.7
Part of a research and development land station	2	3.2
Upstairs of a duplex	1	1.6
Unit built for sale, not condominium	1	1.6
Unit part of apartment complex	1	1.6
Single family detached	4	6.5
No reason	22	35.5

Source: Hartnett (1985).

Cooperatives

In the 1983 survey, out of 196 units reported as a cooperative in the original interview 19 (9.7 percent) were units verified to be not cooperatives (see table 5.29).

This was a significant improvement over the 1979 survey, which had 31.4 percent misclassification for cooperatives. The added cooperative probe did seem to contribute to respondents classifying cooperatives correctly. However, this didn't always work, as noted in the following examples.

Two examples of respondents misclassifying their housing units revealed that they did not understand the AHS definition of a cooperative. One unit was part of a ''farm cooperative'' that converted to a corporation. Another unit was a multifamily household that ''cooperatively'' shared expenses. All explanations for misclassification are listed in table 5.30.

Limitations

The data presented are unweighted tallies from the reinterview questionnaires that were completed by the regional offices. No adjustments have been made for noninterviews or failure to carry out the verification procedures.

Table 5.29. **Verification Results for Housing Units Originally Classified as Cooperatives, 1983 AHS-National**

Item	Number	Percent
Total (occupied and vacant)	**196**	**100**
Verified as cooperative	156	97.6
Verified as not cooperative	19	9.7
Followup not completed	21	10.7
Owner-occupied	**188**	**100**
Verified as cooperative	154	81.9
Verified as not cooperative	19	10.1
Followup not completed	15	8.0
Vacant for sale	**8**	**100**
Verified as cooperative	2	25
Verified as not cooperative	0	0
Followup not completed	6	75

Source: Hartnett (1985).

Table 5.30. **Reasons Cooperatives Were Misclassified, 1983 AHS-National**

Reason	Number	Percent
Total	**19**	**100**
Unit was condominium	12	63.0
Farm coop	1	5.3
Share expenses	1	5.3
Unit is rented from parents	1	5.3
Field representative checked wrong box	1	5.3
No reason	3	15.8

Source: Hartnett (1985).

These results reflect differences for only those housing units that were originally classified as cooperatives or condominiums. It is believed that false positives are also a major source of the gross differences in reporting for these units. The regular reinterview program included questions on cooperative and condominium status for housing units not originally reported as a cooperative or condominium to provide an estimate of errors in the other direction. The results of this latter effort are not available.

MULTIUNIT STRUCTURES FOLLOWUP TO THE 1984 AHS-MS

AHS field representatives have long reported that apartment dwellers often had little knowledge of the structural characteristics of their building. Fuels, heating equipment, and water supply were some of the affected items. However, AHS procedures disallow the use of proxy respondents except in extraordinary cases. Therefore, the quality of AHS structure-specific data in multiunit structures is tied to the impressions of the building's residents.

For example, Smith (1985) analyzed the 1982 AHS-MS reinterview data and found both owners and renters showed moderate to high levels of inconsistency in reporting main heating equipment.

To evaluate the quality of responses from household respondents in multiunit structures and to investigate the feasibility of interviewing structure respondents, a multiunit structure (MUS) followup program was conducted with the 1984 AHS-MS.

Procedures

The MUS was a followup program to the 1984 AHS-MS. All sample units in multiunit structures or multiunit mobile homes were included in the MUS program. The program created a printout of these cases showing selected AHS data and assigned a unique MUS control number to each AHS sample unit. A count of eligible units by metropolitan area is shown below.

Metropolitan area	Number of MUS cases
Birmingham	1,050
Buffalo	1,766
Cleveland	1,589
Indianapolis	1,250
Memphis	1,366
Milwaukee	1,955
Norfolk-Virginia Beach-Newport News	1,296
Oklahoma City	1,269
Providence-Pawtucket-Warwick	1,957
Salt Lake City	1,463
San Jose	1,440
Total	**16,401**

Clerks in the regional offices prepared an AHS-600, Multiunit Structures questionnaire for each address listed. However, they transcribed only the basic address for each case. The Census Bureau suppressed the unit address in order to preserve the confidentiality of the original AHS-MS respondent at the sample unit. This was critical, since a different respondent would be chosen to answer the MUS interview.

The MUS questionnaire was structure-specific rather than unit-specific. It consisted of a subset of the items from the regular AHS-MS questionnaire. Data were collected on:

Interviewer estimate of structure type
Units in structure
Structure type
Presence of commercial establishment
Presence of medical establishment
Water source
Number of apartments sharing a well
Water heating fuel
Presence of public sewer
Other type of sewage disposal
Number of apartments sharing a septic tank
Main heating fuel
Main heating equipment
Supplemental heating equipment
Year structure was built

The MUS questionnaire also referenced the AHS-MS control number(s) of the sample unit(s) in the structure.

AHS-MS field representatives (or current survey interviewers, if the former were unavailable) conducted the MUS interviews in February and March, 1985. The AHS interviews had been completed between June and December, 1984. No AHS field representative could administer the MUS followup in structures where they had obtained the original AHS data. This followup was intended to be done by personal visit. For a small number of cases, regional office personnel conducted telephone interviews with respondents who lived outside the metropolitan area or who requested a telephone interview.

The key to the MUS followup was the definition of an eligible respondent. Because this program attempted to establish ''truth'' about structural systems, only persons with some knowledge of the building were to be chosen. The list of eligible respondents for the MUS followup included: the building's owner, its manager, landlord/landlady, or janitor, the rental or real estate agent for the structure, an official of the condominium or cooperative association (where applicable), or other representatives of the owner or management. Building occupants did not qualify unless they also fit one of the above mentioned categories.

Field representatives were instructed to pick the most knowledgeable respondent when more than one eligible respondent was available.

The completed MUS questionnaires underwent the same clerical edit used for AHS-MS documents. Computer edits were applied to the data, chiefly to resolve problems in matching the AHS and MUS control numbers.

Comparison of AHS and MUS Data

Sample units with different respondents for the AHS and MUS interviews form the universe for comparison. This includes the 10,071 MUS interviews plus 3,069 MUS noninterviews which are ''Associated with another MUS case.'' The followup data for such cases are obtained from the ''master'' MUS interview for the building. Units which are MUS noninterviews because the AHS respondent was an eligible MUS respondent are not included, since we presumed an MUS respondent would be the same person as the AHS respondent.

The AHS and MUS responses were compared using net difference rates. The results are given in table 5.31. For each data item and answer category, the percent of total AHS responses that fell into that cell was linked to the percent of total MUS responses for the cell. (This procedure excludes cases where the item was a nonresponse in either the AHS or MUS interview.) The difference between the two figures is the net difference rate.

This statistic may be interpreted quite simply. The magnitude of the net difference rate reflects the amount of variation between the answers of the AHS and MUS respondents. However, if the MUS respondent is assumed to be a more knowledgeable source than the AHS respondent, the net difference rate also will show the amount and direction of bias in AHS data. A positive rate indicates the AHS respondent overreported the characteristic, while a negative figure indicates underreporting.

The intent of the MUS followup was to identify and interview individuals familiar with the ''truth'' about the AHS unit's housing characteristics. However, a sizable minority of MUS respondents did not fall into the preselected categories of knowledgeable sources (that is, owner, landlord, rental or real estate agent, or condominium or cooperative association official). A further caveat is that some MUS respondents may provide false answers to avoid reporting to the government that they own/manage substandard units, such as housing without indoor plumbing or central heating.

Although the MUS answers will not always be correct, it is assumed that these data are generally better than the AHS responses and that statistically significant difference rates do indicate bias. Certainly the majority of the MUS respondents hold positions allowing them information about the structures, and the field representatives' training emphasized the importance of the respondent for this program. It should follow then that the field representatives chose respondents with some care.

Table 5.31. **Net Difference Rates for Selected Housing Characteristics Estimated From the Multiunit Structures Followup, 11 Metropolitan Areas: 1984 AHS-MS**

Data item	Percent in class		Net difference rate	95-percent confidence interval for net difference rate
	AHS	MUS		
Units in structure (12,168 cases)				
1	0.1	0.1	-	(NA)
2 to 4	43.8	43.5	+0.4	-0.1 to 0.9
5 to 9	23.0	21.1	[1]+1.9	1.3 to 2.5
10 to 19	15.7	15.6	+0.1	-0.5 to 0.6
20 to 49	9.0	9.1	-0.1	-0.5 to 0.3
50 to 99	3.2	4.2	[1]-1.0	-1.3 to -0.7
100 or more	5.2	6.5	[1]-1.3	-1.6 to -1.1
Main heating equipment (12,576 cases)				
Warm-air furnace	51.4	48.9	[1]+2.5	1.8 to 3.2
Steam/hot water system	21.1	24.8	[1]-3.7	-4.2 to -3.1
Heat pump	1.4	1.6	-0.2	-0.5 to +0.0
Built-in electric	8.9	9.2	-0.2	-0.7 to 0.2
Floor/wall furnace	8.0	8.0	+0.1	-0.4 to 0.5
Vented room heaters	5.8	5.4	+0.4	-0.0 to 0.8
Unvented room heaters	1.6	0.9	[1]+0.7	0.5 to 0.9
Portable electric heaters	0.3	0.0	[1]+0.3	0.2 to 0.4
Stoves	1.0	0.5	[1]+0.5	0.3 to 0.7
Fireplaces with inserts	0.0	0.0	-	(NA)
Fireplaces without inserts	0.0	0.0	-	(NA)
Other	0.2	0.7	[1]-0.5	-0.6 to -0.3
None	0.2	0.0	-	(NA)
Main heating fuel (12,467 cases)				
Electricity	28.4	25.1	[1]+3.3	2.8 to 3.8
Gas	64.0	68.0	[1]-4.0	-4.6 to -3.4
Fuel oil	5.8	5.8	+0.0	-0.3 to 0.3
Kerosene	0.2	0.1	-	(NA)
Coal or coke	0.1	0.1	-	(NA)
Wood	0.3	0.1	[1]+0.2	0.1 to 0.3
Solar	0.0	0.0	-	(NA)
Other	1.0	0.9	+0.2	-0.1 to 0.4
None	0.2	0.0	-	(NA)
Water supply (12,635 cases)				
Public system	98.9	99.0	-0.1	-0.2 to 0.1
Well	1.0	1.0	-0.0	-0.1 to 0.1
Spring	0.0	0.0	-	(NA)
Cistern	0.0	0.0	-	(NA)
Other	0.1	0.0	-	(NA)
Public sewer (12,571 cases)				
Yes	97.7	97.8	-0.1	-0.3 to 0.1
No	2.3	2.2	0.1	-0.1 to 0.3
Water Heating fuel (12,010 cases)				
Electricity	28.5	24.1	[1]+4.4	3.8 to 5.0
Gas	66.8	71.5	[1]-4.7	-5.4 to -4.1
Fuel oil	3.9	3.8	+0.1	-0.2 to 0.4
Water Heating fuel (12,010 cases)—Con.				
Kerosene	0.0	0.0	-	(NA)
Coal or coke	0.0	0.0	-	(NA)
Wood	0.1	0.0	-	(NA)
Solar	0.3	0.3	+0.0	-0.1 to 0.1
Other	0.4	0.3	0.1	-0.1 to 0.2
Commercial establishment on property (352 cases)				
Yes	3.1	4.5	-	(NA)
No	96.9	95.5	+1.4	-0.9 to 3.7
Medical establishment on property (355 cases)				
Yes	1.1	0.8	-	(NA)
No	98.9	99.2	-0.3	-1.9 to 1.3
Sewage disposal for units without sewers (199 cases)				
Septic tank/cesspool	100.0	100.0	0.0	-2.0 to 2.0
Outhouse	0.0	0.0	-	(NA)
Other	0.0	0.0	-	(NA)
None	0.0	0.0	-	(NA)
Number of units sharing a septic tank (200 cases)				
Only one	14.0	7.5	[1]+6.5	1.2 to 11.8
2 to 5	70.5	77.0	[1]-6.5	-12.8 to -0.2
6 or more	15.5	15.5	0.0	-3.9 to 3.9
Number of units sharing a well (89 cases)				
Only 1	13.5	12.4	-	(NA)
2 to 5	68.5	73.0	-4.5	-16.2 to 7.2
6 or more	18.0	14.6	-	(NA)
Year structure built (11,164 cases)				
1984	1.3	1.5	[1]-0.2	-0.3 to -0.1
1983	0.8	0.8	-0.1	-0.2 to 0.1
1982	0.9	0.9	+0.1	-0.1 to 0.2
1981	1.0	0.9	+0.1	-0.1 to 0.3
1980	1.0	0.9	+0.1	-0.2 to 0.3
1979	2.1	2.0	+0.1	-0.2 to 0.4
1975 to 1978	7.5	7.9	-0.4	-0.9 to 0.1
1970 to 1974	19.8	19.8	-0.0	-0.8 to 0.7
1960 to 1969	20.2	21.7	[1]-1.5	-2.2 to -0.7
1950 to 1959	9.0	8.0	[1]+1.0	0.4 to 1.7
1940 to 1949	7.0	6.5	+0.5	-0.1 to 1.0
1930 to 1939	6.7	5.8	[1]+0.8	0.3 to 1.4
1920 to 1929	7.9	8.2	-0.4	-1.0 to 0.3
1919 or earlier	14.9	15.1	-0.2	-0.8 to 0.4

(NA) Not applicable.
* Indicates net difference rate is significant at the 5-percent level.
– Indicates net difference rate not shown where category contains fewer than 40 cases.

Source: Williams (1985).

It can be seen that considerable variation exists in the rate at which the AHS and MUS responses matched for the 12 items analyzed. Several questions show no significant difference between the two sets of responses for all answer categories on which net difference rates could be calculated. These included: water supply, presence of public sewer, commercial establishment on property, medical establishment on property, sewage disposal for units without sewers, and number of units sharing a well. On the other hand, data items such as units in structure, main heating equipment, main heating fuel, and number of units sharing a septic tank, had a significant difference on majority of the answer categories.

The patterns in the quality of the AHS data are apparent. Heating equipment and fuels tend to be more poorly reported. Not only do several of their answer categories show bias, but the categories involved are those most frequently reported for the characteristics. Electricity is overreported as a home-heating and water-heating fuel, while gas is underreported in AHS on the same two items. A similar process occurs with the warm-air furnace (overreported) and the steam or hot water system (underreported) categories of main heating equipment. Apartment dwellers' ignorance of these items is not surprising. They are frequently physically separated from their heat sources, nor do they have responsibility for maintenance in most cases.

The data show that AHS respondents overreported unvented room heaters, portable electric heaters, and stoves as main heating equipment. Some of the difference may be in fact the structure respondent's underreporting of these systems. As previously mentioned, building owners may not care to admit poor housing quality to representatives of the Federal Government.

Generally, the questions dealing with water and sewage had good agreement between the AHS and MUS respondents. In these multiunit structures, nearly all units will have public water (99.0 percent) and sewers (97.8 percent). Those without sewers have septic tanks without fail (100 percent). This means AHS respondents who can only guess these characteristics are quite safe choosing the familiar answer.

The data concerning the number of units sharing a well and the number of units sharing a septic tank are very similar in the type of information elicited. It is curious then that the answers for the latter, but not the former, are significantly different between the AHS and MUS respondents. Perhaps the key lies in the fact that the first item had less than half the good responses of the second, requiring, therefore, a comparatively high net difference rate in order for the data to be statistically different.

The variables commercial establishment on property and medical establishment on property demonstrate good agreement between the AHS and MUS replies. However the AHS questionnaire limits these questions to owner-occupied units, a small subset (about 15 percent) of the occupants of multiunit structures. The AHS respondents for these items also may be more knowledgeable than AHS respondents in general.

Units-in-structure shows bias in three of its six answer categories. The unit respondents reported too many 5-to-9-unit structures and too few of the 50-or-more-unit buildings. Most times when the AHS respondent misreported units-in-structure for 5-to-9 or 50-to-99-unit buildings, they were off by only one answer category. However for the 100-or-more-unit structures, a little over three-fourths (76.7 percent) of the wrong answers were at least two categories removed from the MUS response. So for these cases, the AHS respondents miscounted the buildings by at least 50 apartments. The published data will palliate the effects of the problem somewhat because the upper tail category for units-in-structure will be 50-or-more units.

The probable explanation for the poor showing of units-in-structure data for very large structures is that townhouse and/or garden-type apartment buildings are involved. The problem which would then confront both AHS and MUS respondents, would be to determine where the dividing line between separate structures occurred.

These data underscore the fact that respondents do not have a clear idea of the definition of a ''building.'' This confusion persists in the AHS data even after a special set of questions were added in 1984 to clarify the matter. Since similar comparative data are not available for the AHS prior to 1984, the amount of amelioration by the new items is unknown.

The analysis of the year built data is instructive. Only 4 of 14 answer categories have significant differences between the rates of AHS and MUS responses. These categories include those for the current survey year and for the decades of the sixties, fifties, and thirties. The first two were underreported by AHS respondents, the latter two had positive biases. The recent and distant past did not suffer these distortions.

Among the four biased categories, the majority of the erring AHS respondents did report year built within one answer category (plus or minus) of the MUS interval. In fact, over 80 percent of the AHS replies which did not match the MUS response of ''1960-69'' were either ''1970-74'' or ''1950-59.'' Only 51 percent of the similar statistics for the MUS interval ''1930-39'' was either ''1940-49'' or ''1920-29.'' It appears that the AHS respondents have a good general idea of their building's age. However, the answer intervals are already so broad (up to 10 years, excluding the earliest category) that ''near misses'' are not very near.

Several factors may influence the quality of the AHS data. An obvious concern is the source of the information. Table 5.32 provides a description of AHS participants in multiunit structures. The first two categories, plus ''proxy respondent,'' are respondents for occupied units. Respondents for vacant units are identified by title. Therefore, the category ''owner'' includes only cases where a nonresident owner provided information about a vacant unit.

Table 5.32. **Distribution of 1984 AHS-MS Respondents for Cases Included in the MUS Followup**

Type of AHS respondent	Total cases	Percent of cases
First occupant listed in the control card	9,751	74.2
Other household member	2,565	19.5
Owner	62	0.5
Landlord	39	0.3
Rental agent	103	0.8
Neighbor	211	1.6
Field representative observation	10	0.1
Proxy respondent	7	0.1
Other	220	1.7
Not reported	170	1.3
Out of range	2	0.0

Source: Williams (1985).

'Line 1'' denotes the first occupant listed on the AHS control card, almost always the person who owns or rents the sample unit. In the category, ''other household member,'' 92.4 percent of these respondents were listed on line 2 of the control card. Presumably, many of these people are spouses of the owner or renter. The remaining 7.6 percent of ''other household members'' were listed on lines 3 through 11 of the AHS control card. The ''out of range'' category is made up of two cases with a code of 29 which may be a keying transposition of 92, the code for rental agent.

It is possible to examine net difference rates by respondent type, however, there are some important limitations. Due to the small number of cases for some of the respondent classifications, net difference rates cannot be shown for the landlord, field representative observation, proxy respondent, and out-of-range groups. (To provide information about these types of respondents, aggregated tables were compiled for all respondents from vacant interviews.) Within data items, a similar problem with data reliability occurs because, for example, most of the responses fall into 2 or 3 categories of a 12 category distribution. Thus many answer categories have too few cases on which to calculate net difference rates. Household respondents, particularly line 1, make the poorest showing. Heating equipment, heating fuels (both home- and water-heating), year built and units-in-structure suffered in accuracy (in decreasing order of severity) at the hands of these individuals. Few differences appear between the line 1 respondents and the total respondents regarding which characteristics showed bias and the direction of the bias. However, in contrast to the overall net difference rates, line 1 respondents produced no bias on the item of number of units sharing a septic tank. This subset of respondents overreported units built in 1979, but failed to show any bias for units built in the fifties and thirties. It should be noted that five of the answer categories that were biased for all AHS were not examined for these respondents due to the small universe of replies for the latter.

In general, persons providing information for vacant units demonstrated less bias than household respondents. This result may reflect the fact that field representatives must seek out knowledgeable respondents for vacant units. For occupied units, only adult occupants may provide the AHS information regardless of their degree of familiarity with the unit. (Although our preference is the most knowledgeable adult occupant.)

At vacant units, the data items—units-in-structure, water heating fuel, and number of units sharing a septic tank—had no biased answer categories, unlike the same data for all cases. Only for a few categories of the main heating equipment, main heating fuel, and year built items was any bias displayed. The category, ''vented room heaters'' had a positive bias for vacant units, but no bias for occupied units. Neighbors as respondents seem to be the source of this error. The year built category, 1984, is overreported by vacant unit respondents, and underreported by household respondents. Otherwise the two remaining vacant interview items that showed bias share this fact and the direction of the bias with the data from household respondents.

In addition to the type of respondent, other factors may influence AHS data quality from multiunit structures. Units-in-structure is an obvious candidate. Occupants of large buildings may be less informed about the structure than those in smaller buildings. The analysis of the net difference rate data by units-in-structure showed that as the size of the sample structure changed so did the identity and direction of data categories exhibiting bias.

For the variable main heating equipment, steam or hot water systems were underreported in all structure sizes. However, warm-air furnaces were overreported only in buildings with at least 10 units. Two categories, built-in electric heaters and floor, wall, or pipeless furnaces were underreported in 100-or-more-unit structures, but nowhere else. At the other end of the scale, bias in reporting stoves and ''other'' as main heating equipment was seen only in 2-to-4-unit structures. Bias appeared in the categories vented and unvented room heaters, only in small-to-mid-sized buildings. Portable electric heaters which are overreported for the total cases did not show bias in any of the units-in-structure subgroups. This was due to the fact that the small number of cases in the category could not generate net difference rates when spread over several units-in-structure groups.

The data for main heating fuel present simple picture. As with the total cases, electricity is overreported in each units-in-structure category while gas is universally underreported. Fuel oil has a positive bias in 100-or-more-unit buildings, but is unbiased elsewhere. In contrast, the ''other'' fuel category is biased only in lower-sized structures (nine or fewer units). Wood as a main fuel is overreported for total cases, but like portable electric heaters, the number of sample cases for wood is too small to show statistics by units-in-structure.

Within the data items—water supply, presence of public sewer, commercial establishment on property, medical establishment on property, sewage disposal for units without sewers, and number of units sharing a well—only one

data item and one structure size category deviates from the pattern of the total cases. In the total cases, no bias occurred in all of the answer categories for which net difference rates could be calculated. When the data are grouped by units-in-structure, these data show that 5-to-9-unit buildings underreport the presence of public sewers.

The variable number of units sharing a septic tank shows bias in two of its answer categories for total units. Only one net difference rate, that for 2-to-4-unit structures reporting 2-to-5 units sharing the septic tank, could be produced for the units-in-structure groupings. This statistic matched the direction of bias of the total cases.

The smaller structures (nine-or-fewer units) have proportionately fewer biased year-built categories than the larger buildings.

Regardless of the structure's size, AHS respondents tended to give biased answers for the same four data items: heating equipment, home- and water-heating fuels, and year built.

The MUS followup was a one-time operation. As noted in Williams (1985) the MUS was relatively expensive for the amount of data improvement that resulted. Based on the results of the MUS followup, the AHS questionnaire items related to heating equipment were changed to improve the reporting for this item. There are no current plans to supplement or replace AHS household respondents' information with data from other sources.

Chapter 6.
Data Processing

OVERVIEW OF DATA PROCESSING PROCEDURES

Processing of the data is an integral part of the AHS survey and its proper operation has a large effect on the accuracy of the data. Data processing procedures for AHS-National and MS are essentially same. The data preparation for AHS has two main phases. The first phase takes place in the Census Bureau's regional offices. Clerical personnel edit a sample of the completed questionnaires received from the AHS field representatives. Data entry clerks key the information from these edited and the unedited questionnaires. The resulting data files are transmitted electronically to Census Bureau headquarters.

The second phase of data preparation consists of a receipt and control operation to ensure that all assigned sample cases have been accounted for, and a series of computer runs to edit for consistency and impute missing values. These operations are carried out at headquarters.

EDITING

The first operation in the regional offices is a clerical edit of completed questionnaires mailed in by each field representative. This check detects omissions and other errors in the completion of the questionnaires. For new field representatives, the questionnaires from their first assignment are fully edited. If their work is satisfactory, the clerks edit only four or five questionnaires from each subsequent assignment.

The next step is data entry: keying information from control cards and questionnaires. Edits are built into the data entry program to ensure that:

1. The data are keyed in the proper sequence.
2. Certain key identifiers, such as control number, name, and relationship to householder, are present.
3. Selected numeric items, mostly on the control card, are present.

Data failing these edits are rekeyed after investigation and correction. Data files for the accepted batches are transmitted electronically to headquarters.

The initial step with files received from the regional offices is a receipt and control run to ensure that all expected cases, whether interviews or noninterviews, are received. Errors identified in this step are described in ''reject listings'' for the regional offices. Regional office personnel resolve the problems by reviewing the completed questionnaires or contacting field representatives. Corrections and additional data are keyed and transmitted to headquarters.

Subsequent steps in data preparation are:

1. Data are imputed for selected missing items in interviewed units. Imputation also is used to replace reported values that fail consistency tests in editing. The Census Bureau's traditional ''hot-deck'' procedure (Bailar et al., 1978) is used for imputation. The variables used to define imputation matrices depend on the item being imputed and generally vary widely from item to item.
2. An edit is performed to ensure consistency of responses recorded for units, persons, families, and households. Consistency is examined within and between sections of the questionnaire and between the control card and the questionnaire.
3. Each section of the questionnaire is edited to ensure that responses appear where they should.
4. Recodes based on combinations of data items are added to the records and the codes that identify geographical areas are corrected if necessary. Confidential name and address information is removed from the file.

At this point the data are ready for weighting and estimation, as described in chapter 7.

QUALITY CONTROL OPERATIONS IN DATA PROCESSING

Clerical Edit

The Regional Offices (RO's) clerically edit a sample of each field representative's (FR's) work to check for errors. The RO's send the results of this check to the FR's, along with instructions, if necessary, on correcting errors found in the clerical edit. The RO's give each FR a rating based on the results of this edit.

Data Keying

The work of each new data keyer is verified (rekeyed as a check for errors) 100 percent. After the keyer's first several batches of documents are verified, the RO calculates an error rate. If the error rate is at or below the

acceptable standard for a keyer (0.4 percent errors or fewer), then the keyer's work is checked (verified) using a sample. If the keyer's error rate rises above the acceptable standard, then the keyer's work is verified 100 percent until the error rate drops to the acceptable standard.

The data keying programs that AHS uses also have edits to check the quality of the keying. These edits check for an acceptable range of entries in each item and the appropriate parts of the questionnaires are keyed. Questionnaires that fail selected data keying edit checks cannot be transmitted until the RO corrects the error.

Pre-edit

After the data are keyed and transmitted, they are run through a computer program that checks for acceptable entries in selected items. If there is a problem, the record is ''rejected'' and the RO's review the situation clerically and make appropriate correction

Computer Edit

The Demographic Surveys Division (DSD) at Census Bureau headquarters runs the computer edit for AHS using division-developed software called Record/Item Management (RIM). This software does an automatic range check of the entries in each item. If there is an entry that is not within the acceptable range, the system flags the entry. Also, the system uses a data dictionary to give names to each data variable. This helps in coding the programs because the programmers can refer to a data item by name instead of referring to a place on a record layout, thereby reducing coding errors.

QUALITY ASSURANCE RESULTS FOR KEYING 1989 AHS-NATIONAL

Quality control operations in data keying were outlined in the section ''Quality Control Operations in Data Processing,'' in this chapter, page 67. In this section, we describe the quality control operations methodology.

Methodology

AHS-National batches for verification of data keying fall into three categories: noninterview, vacant, and interview. Noninterview and vacant batches consist of 15 control cards and questionnaires, and are always 100-percent verified.

Interview batches contain 15 control cards and their associated questionnaires. Batches keyed by ''qualified keyers'' are verified on a sample basis. One-fifth of the forms in a batch fall into sample verification, with control cards and questionnaires counted separately. Batches keyed by ''nonqualified keyers'' are verified 100 percent.

If a qualified keyer's sample verification error rate exceeds the acceptable level for a calendar month, the keyer must requalify (by keying on 100-percent verification) the next calendar month. A keyer's sample verification error rate is ''acceptable'' if his/her sample error rate is within three standard deviations of the national sample verification process average error rate.

The following QC operations are automated:

1. Sampling
2. Detecting differences between keyer and verifier
3. Tallying the errors

The targeted average outgoing quality limit (AOQL) for AHS-National keying is 0.40 percent. That is, we desire that the outgoing error rate with regard to keying, after all inspection steps are performed, to be at most 0.40 percent. This target of 0.40 percent is used to come up with a set of acceptance criteria which are applied to each batch of work subject to Quality Control (QC). The specifications for QC are documented by Wetzel (1990).

Error Rates and Rejection Rates

Keying under sample verification, 100-percent verification (interview), and 100-percent verification noninterview and vacant represent four different processes, each with different quality levels. The field error rates and the batch reject rates for each category are reported separately. For 100-percent verified batches, the rejection rate is the proportion of batches with unacceptable error rates. By unacceptable we mean that if the batch in question had been sample verified, it would have been rejected.

Results of AHS-National Keying Verification

Table 6.1 summarizes the national results for the four types of keying verification.

Field error rates and batch reject rates from 100-percent verification are higher than those from sample verification. We expect higher rates from 100-percent verification because ''unqualified keyers'' usually perform the work.

Table 6.1. **National Results—All Types of Keying, 1989 AHS-National**

Type of keying	Batches keyed	Incoming error rate (percent)	Rejection rate (percent)
Sample-verified interview	1973	0.16	6
100-percent verified interview	442	0.82	79
100-percent verified noninterview	327	0.29	14
100-percent verified vacant	349	0.32	37

Source: Waite (1990f).

Some possible causes of high error rates include:

1. Poor keyer training
2. Inability to hire good keyers
3. Too much turnover of the keying staff
4. Problems with the equipment or facilities

The national average incoming sample verification error rate (before correction of errors—rectification) for AHS-National keying was 0.16 percent. Note that this was lower than the specified AOQL for keying of 0.40 percent.

QUALITY ASSURANCE RESULTS FOR KEYING 1989 AHS-MS

The methodology for quality control operations in data keying for AHS-MS is the same as the methodology for AHS-National described in the section ''Quality Assurance Results for Keying 1989 AHS-National,'' in this chapter, page 68. In this section, we provide results for keying 1989 AHS-MS data from Waite (1990d).

Table 6.2 is a national summary of the error and rejection rates for the various types of AHS-MS keying. The national average error rate is higher for 100-percent verified interview batches than for sample verified interview batches because only qualified keyers are to be sample verified.

Noninterview and vacant batches include the AHS-63 questionnaire, a different form than the regular interview batches. If the batches are distributed at random to keyers for initial keying, some keyers mightnever receive enough experience with the AHS-63 to key as accurately as the regular batches. By assigning these batch types to specific keyers, we believe the error rates are reduced through experience.

Rectification

Rectification is the process of correcting errors in reject batches. If a batch is rejected, all previously nonverified documents are verified and corrected. All rejected batches must be rectified to maintain the 0.40 percent AOQL. It is important that the RO's use the right kind of verification to maintain quality and keep down costs.

Table 6.2. **National Results—All Types of Keying, 1989 AHS-MS**

Type of keying	Batches keyed	Error rate (percent	Rejection rate (percent)
Sample-verified interview	1775	0.19	6
100-percent verified interview	314	0.71	74
100-percent verified noninterview .	411	0.36	18
100-percent verified vacant	282	0.29	21

Source: Waite (1990d).

RESULTS OF RESEARCH ON REGIONAL OFFICE PRE-EDIT FOR 1989 AHS-MS

The regional office pre-edit is designed to improve the quality of the survey data. Data records (information as keyed from the AHS-61 Control Card and the AHS-62/AHS-63 questionnaires) are rejected if they fail to meet certain standards. Regional Office staff research the problems causing the records to reject, enter the corrective actions needed on the Correction Section of the Reject Listing, and key these corrections. For all metropolitan areas, the 1989 AHS-MS Regional Office pre-edit was conducted in four waves. The first wave was run after the keying was completed for panel 04 (panels 04 and 05 for Detroit). The second and subsequent waves included additional panels, as well as panels that had already been included in earlier wave(s). Rejected records that were not resolved in the earlier wave(s) were rejected again. Chapter 20 of the 1989 AHS-MS Office Manual provides an overview of the pre-edit operations and detailed instructions to the Regional Office for researching and processing the pre-edit rejects.

Abernathy (1991) analyzes wave 1 reject data to (1) determine the status of the rejects, (2) determine the types of errors that caused the records to reject, and (3) compare the pre-edit reject corrections with how the reject situations would have been edited during the computer edit. The results are summarized in this section.

Status of Reject

There were 2,784 records that were rejected for 52 reject reasons. The ''Status of Reject'' research showed that:

- Eighty-three percent of the total rejects were resolved. The RO's used the ''Accept'' command for another 7 percent.
- The rejected data were resolved and accepted for 15 percent of the cases that were rejected for ''type of living quarters inconsistent,'' 64 percent of the cases for ''units in structure inconsistent,'' and 41 percent of the cases for ''year built inconsistent.''
- Five percent of the total rejects were *not* resolved even after a correction was made and rejected again in Wave 2.
- No correction was made for 4 percent of the total rejects, some of which rejected again in Wave 2.

Based on the above percentages, almost all of the rejects were either resolved or the keyed entry accepted. A small percent of the cases were rejected again in Wave 2 for the same reject reason. The percent of records rejected again in Wave 2 does not include records that were resolved in Wave 1 for a specific reject, but rejected again in Wave 2 for a different reject reason despite the Wave 1

correction. The research concentrated only on the status of the Wave 1 rejects. Abernathy suggests that further research be done on ''previous rejects,'' that is, records rejected both in Wave 1 and Wave 2, to get a better feel of the incidence of records being rejected again for different reasons.

Type of Error

The ''Type of Error'' research showed that:

- Fifteen percent of the total rejects were caused by relationship code errors.
- Seventy-seven percent of the total rejects were because of specific data errors.
- Eight percent of the total rejects were because of other errors.

For cases rejecting as ''nonrelative code missing,'' ''reference person has illegal relatives,'' ''reference person has two spouses,'' ''husband not married male,'' and ''wife not married female,'' various relationship code errors or omissions were present. However, for ''reference person error,'' more than half of the relationship errors were due to missing relationship code(s), and for ''reference person relatives missing'' the majority of the relationship errors were due to entering the incorrect reference person relationship code (''1''–reference person with relatives—entered instead of ''2''–reference person without relatives).

The specific data errors for many of the reject reasons are due to not properly editing the survey forms. This is especially true for the following reject reasons where all, or most, of the specific data errors are due to omissions:

Reject reason	Omission
Relationship code missing	Control card item 13
Illegal spouse for husband/wife	Control card item 23
Illegal parent of child	Control card item 16
Control number status missing	Control card item 6
Income line number missing	AHS-62 item 114
Nonrelative missing	AHS-62 item 184

For ''illegal age,'' most of the specific data errors were due to entering an incorrect age, therefore suggesting that the Age Verification Chart (included in the flashcard booklet of the field representatives) is not being used.

The records that rejected for ''units in structure inconsistent'' (with prior year data) were included in the ''Specific Data'' category because of the nature of the reject. Earlier research conducted by Abernathy (1987) on this reject reason showed that the units in structure entry for some one-unit attached and multiunit structures had a tendency to change from one period to the next (see the section ''1985 AHS-MS Reinterview,'' chapter 5, page 54).

The main reason for the specific data errors for ''mover missing'' was not apparent. The errors were split between situations where the mover line numbers and some other data were present (except Zone Code) *and* situations where mover data (including line numbers) were completely blank. Abernathy suggests that further research be done for ''mover missing'' to determine: 1) if by revising AHS-62 items 51a-b beginning with the 1990 MS, the number of completely blank mover columns changed, and 2) how often correction data other than ''blank'' was entered for Zone Code.

For ''year built inconsistent,'' all the rejects were included as ''specific data'' errors, even though in most cases the year built entry was probably the one provided by the respondent.

Computer Edit Action for Item/Source Code

The comparison of the computer edit and pre-edit correction for all RO's showed that:

- Some reject situations (for ''type of living quarters inconsistent,'' ''illegal age of parent,'' ''illegal age,'' and ''units in structure inconsistent'') are not specifically addressed in the computer edits. This means that unless the entry is inconsistent with the entry in another item, the entry was accepted as keyed on the final edited file.
- The computer edit action was the same as the pre-edit action for fewer than half (45 percent) of the reject situations. However, for household demographic characteristics about 60 percent of the correction actions were the same as those the computer edits would have applied for these reject reasons.
- The computer edit action was different from the pre-edit action for 35 percent of the reject situations.
- The computer edit action could not be determined for about 5 percent of the reject situations.

Even though the overall percentages were 45 percent for ''same'' action and 35 percent for ''different'' action, the distribution by reject reason varied. For some reject reasons (for example, ''type of living quarters missing,'' ''reference person error,'' ''nonrelative code missing,'' ''reference person relatives missing,'' ''reference person has illegal relatives,'' ''illegal spouse for husband/wife,'' and ''illegal parent of child'') most of the actions taken in the pre-edit were *the same* as the computer edit action. Whereas, for some other reject reasons (for example, ''control number status missing,'' ''nonrelative rent missing,'' and ''nonrelative missing'') most of the actions taken in the pre-edit were *different* from the computer edit action. ''Mover missing'' and ''year built inconsistent'' both had a large percentage of ''different'' as well as ''DK Computer Action.'' For ''mover missing,'' 22 of the different actions were because the RO's entered ''blank'' as the pre-edit

correction for Zone Code and/or geography code. Most of the remaining 21 cases were because a zone code of 00 (Off Map) but neither a blank nor a geography code was entered for source code 2350—geography code. For ‘‘year built inconsistent,’’ the different actions were either because the ‘‘accept’’ command was used or a blank entry was entered.

Chapter 7.
Weighting the Sample

INTRODUCTION

Weighting is necessary to convert raw data from AHS into statistics that can be used for descriptive and analytical purposes. These procedures have three goals: to minimize biases that may result from unit and item nonresponse; to take account of the selection probabilities used at every stage of sample selection; and to make use of data from external sources, such as the decennial census, to improve the precision of AHS estimates. Estimation procedures for AHS-National and AHS-MS are different because sample designs are different.

ESTIMATION FOR AHS-NATIONAL

In the 1989 AHS-National, the final weight used in tabulating housing inventory characteristics was equal to the following product:

(base weight) x (duplication control factor)
x (permit new construction noninterview adjustment factor)
x (type A unable-to-locate factor)
x (type A noninterview adjustment factor)
x (first-stage ratio estimate factor)
x (second-stage ratio estimate factor)
x (third-stage ratio estimate factor)
x (ratio estimate factor from all rankings)

A brief explanation for each component is given below. A detailed specification for weighting is given in Waite (1990a).

Base weight. The base weight is the reciprocal of the probability of selection for a given sample unit.

Duplication control factor (DCF). A duplication control factor (DCF) is used to adjust the basic weight of a unit to reflect the correct probability of selection. There are three situations for which DCF's are computed.

1. During the listing of a structure too many (more than 15) units are found and a subsample is selected to make field representatives' workloads manageable. For example, if a structure has 20 units and a subsample of 10 units is selected for interview, these 10 units are given a DCF of 2.000.
2. The designated number of units could not be selected from a segment. For example, if only three units can be selected instead of the designated four units from a segment, a DCF of 1.333 is applied to the three units selected.
3. All within-structure changes in multiunit structures are selected unless there are too many units. For example, a field representative finds two added units in a structure that had 10 units at the time of original sample selection. The new units receive a DCF of 0.1 because the new units would have been found if any one of the 10 original units in the structure were in sample.

In some cases, a unit can receive more than one DCF. All DCF's for a unit are multiplied to derive the final DCF for the unit. The maximum DCF used for a unit is 4, as a tradeoff between unbiased estimation and variance reduction. All records not requiring a DCF are given an implied DCF of 1.000.

The weight = (base weight) x (duplication control factor) reflects the correct probability of selection for each unit. The remaining steps in the weighting consist of two phases. In the first phase, a series of adjustments are made to account for units which could not be interviewed for a number of reasons.

Permit new construction noninterview adjustment factor. The permit new construction (NC) noninterview adjustment factor is calculated, using permit segments only, to account for units for which permits are unavailable for sampling and units which cannot be located. For each region a set of factors is calculated for the cells in table 7.1.

The noninterview adjustment factor F_c for cell c is given by

$$F_c = \frac{I_c + NI_c}{I_c} \qquad (7.1)$$

Table 7.1. **Permit New Construction Noninterview Adjustment Cells and Scale Values**

Inside MSA		Outside MSA
Inside central city	Not in central city	
10	20	40

Where,

I_c = Weighted sum of interviews, type-A noninterviews (except type A unable-to- locate), type-B noninterviews, type-C noninterviews and ineligible vacants in cell c.

NI_c = Weighted sum of type-M and type-A unable-to-locate noninterviews in cell c. A Type M noninterview occurs when a sample is selected from the Building Permits Survey. After that, a matching is done for the permit address listing, if the permit is nonavailable or accesible is assigned a Type M noninterview code. (See the section ''Noninterviews,'' chapter 3, page 29, for the definition of type A, B, and C noninterviews.)

The weight used to obtain the above weighted counts is (base weight) x (duplication control factor). The following conditions must be met before the noninterview factors can be applied to appropriate records:

1. A cell must have at least 30 sample cases and at least one noninterview case.
2. The noninterview factor must be less than 1.5. A cell which fails either of these two criteria is combined with the cell that has the nearest scale value. For example, if the cell ''Inside MSA, not in central city'' in table 7.1 fails to meet the conditions, it would be combined with the cell ''Inside MSA, inside central city.'' The combined cell will have a scale value 15 (that is, (10+20)/2). Collapsing is continued until the conditions are met.

Type-A unable-to-locate adjustment factor. Some of the addresses in address units, and HUCS segments could not be located by the field representatives. This factor adjusts for these unable-to-locate units and is calculated for the cells in table 7.2 separately within each region.

The type-A unable-to-locate adjustment factor F_c has the same form as (7.1) with

I_c = Weighted sum of units with a Census serial number on the master hit tape (this includes all interviews, and type-A, type-B, and type-C noninterviews, but excludes type-A unable-to-locate units) in cell c.

and

NI_c = Weighted sum of type-A unable-to-locate units with census serial numbers in cell c.

Table 7.2. **Type-A Unable-to-Locate Adjustment Cells**

	Inside MSA		
	In central city	Not in central city	Outside MSA
Address segments . . .			
Unit segments			
HUCS segments			

The weight used in calculating this factor is (base weight) x (duplication control factor).

Type-A noninterview adjustment factor. The Type A noninterview adjustment accounts for units which could not be interviewed because either no one was home after repeated visits or the respondent refused to be interviewed. When 1987 or 1985 AHS or 1980 census data were available, this information was used to determine the noninterview adjustment cell to which the unit belongs. The cells are defined by characteristics such as tenure, geography, number of units in structure, and number of rooms on the basis of the noninterview adjustment research documented by Parmer (1986). When previous data were not available, adjustment factors were computed separately using more general characteristics such as type of area and type of housing unit (that is, mobile home, nonmobile home). The adjustment factor has the same form as in formula 7.1, page 77. A detailed description of this adjustment is given in Waite (1990a).

The second phase of estimation involves a three-stage ratio adjustment procedure to account for the sampling of nonself-representing PSU's, to account for known sampling deficiencies in new construction, and to bring the sample estimate of housing units into close agreement with estimates derived from independent sources for several key characteristics.

First-Stage Ratio Estimation Factor

The first-stage of ratio adjustment is employed to reduce the component of variance due to sampling of nonself-representing PSU's. The procedure takes into account the differences that existed at the time of the 1980 census between the number of housing units estimated from the nonself-representing sample PSU's and the actual 1980 census count of housing units from all nonself-representing strata. Factors accounting for these differences were computed separately for 15 place-of-residence/tenure cells for the Northeast and Midwest regions, 35 place-of-residence/ethnicity-race/tenure cells for the South region, and 25 place-of-residence/ethnicity/tenure cells for the West region. The first-stage ratio estimation factor is equal to the following ratio:

$$\frac{\text{Actual 1980 census housing units in a cell for all nonself-representing strata}}{\text{Number of 1980 housing units in the same cell estimated from the sample nonself-representing PSU's}}$$

The numerator of the ratio for a cell is calculated by summing the 1980 census housing unit counts for that cell across all nonself-representing strata. The denominator is calculated by weighting the 1980 census housing unit counts from each nonself-representing sample PSU by the inverse of the probability of selection for the PSU and summing the weighted counts across all nonself-representing sample PSU's.

The first-stage ratio adjustment factors for the 1989 AHS-National Survey for the Northeast, Midwest, South, and Western regions are given in tables 7.3, 7.4, 7.5, and 7.6, respectively.

Second-Stage Ratio Estimation Factor

The second stage of the ratio estimation procedure is employed to adjust the AHS sample estimates of old construction (occupied and vacant HU's) and new construction (that is, number of units built since the 1980 census) to account for known deficiencies in the AHS sample. For nonmobile homes, the sample estimates are controlled to independently derived estimates from the Survey of Construction. For mobile homes, the sample estimates are controlled to independently derived estimates from the Survey of Mobile Home Placements. These estimates are considered to be the best estimates available for these types of units. Factors are computed separately for each region. The second-stage factor is equal to the following ratio:

$$\frac{\text{Independently derived estimate for a cell}}{\text{AHS sample estimate in that cell}}$$

Table 7.3. **First-Stage Factors for the Northeast Region: 1989 AHS-National**

	Owner	Renter	Vacant
MSA—central city	.59587	.71085	.79348
Balance MSA—urban	.76707	.78993	.79348
Balance MSA—rural	.84231	.78613	.79348
Outside MSA—urban	1.28771	1.31179	1.53772
Outside MSA—rural	1.02651	.97065	.65764

Source: Waite (1990a).

Table 7.4. **First-Stage Factors for the Midwest Region: 1989 AHS-National**

	Owner	Renter	Vacant
MSA—central city	.96722	.93349	.90245
Balance MSA—urban	1.12648	1.17442	1.09787
Balance MSA—rural	1.00778	.93853	1.04735
Outside MSA—urban	1.00639	.98762	1.11161
Outside MSA—rural	1.01874	.99797	1.02663

Source: Waite (1990a).

Table 7.5. **First-Stage Factors for the South Region: 1989 AHS-National**

	Non-Black non-Hispanic			Black non-Hispanic		Hispanic	
	Owner	Renter	Vacant	Owner	Renter	Owner	Renter
MSA—central city	1.12922	1.11361	1.07512	1.04132	1.04832	1.24172	1.21952
Balance MSA—urban	.88988	.98469	.83419	.94768	1.05592	.88988	.98469
Balance MSA—rural	1.14949	1.20911	1.10893	1.61407	1.61407	1.14949	1.20911
Outside MSA—urban	.96023	.93972	.58200	.93329	.92760	1.17485	1.08767
Outside MSA—rural	.98444	.98267	.95295	.90853	.90942	1.38638	1.16735

Source: Waite (1990a).

Table 7.6. **First-Stage Factors for the West Region: 1989 AHS-National**

	Non-Hispanic			Hispanic	
	Owner	Renter	Vacant	Owner	Renter
MSA—central city	.89699	.92879	.83406	.77724	.88964
Balance MSA—urban	.81404	.70904	.79989	.81404	.70904
Balance MSA—rural	.92183	.92470	1.05637	.92183	.92470
Outside MSA—urban	1.15524	1.26790	1.36865	1.03929	1.22645
Outside MSA—rural	1.0053	1.11913	1.01234	.63165	1.06334

Source: Waite (1990a).

Table 7.7. **Second-Stage Ratio Adjustment Factors: 1989 AHS-National**

	Northeast	Midwest	South	West
Old construction				
Occupied	1.0096	1.0130	1.0417	1.0128
Vacants	1.0273	1.0610	1.0724	0.9697
New construction				
Conventional				
4/80 to 12/84	1.1271	1.2878	1.1582	1.1875
1/85 to 12/87	0.9412	1.0239	0.9931	1.0325
1/88 and later	1.3526	1.5672	1.4788	1.4861
Mobile homes				
1980 to 1982	0.5658	1.1686	1.2455	1.1519
1983 to 1985	0.8771	1.3844	0.9951	1.6036
1986 and later	2.3072	2.2709	1.4888	2.5510

Source: Waite (1990a).

The denominator of this ratio is obtained by summing the existing weight on each record after the first stage ratio estimation over all records in a cell in a region. The second-stage ratio estimation factors used in the 1989 weighting are given in table 7.7.

Third-Stage Ratio Estimation Factor

The third stage of the ratio estimation procedure is employed to adjust the AHS sample estimate of housing units to independently derived current estimates for certain key characteristics. It is believed that these characteristics are highly correlated with other characteristics of interest for the AHS. This stage of the procedure is actually done in two steps for occupied units. During the first step, the sample estimate of occupied housing units is controlled to an independently derived estimate for 12 tenure/ethnicity (that is, Hispanic householder—non-Hispanic householder)/household-status cells for each region. After applying the factor computed in this step to the interviewed occupied units, the new sample estimate of occupied housing units is controlled to an independently derived estimate for 12 tenure/race (that is, Black householder—non-Black householder)/household-status cells for each region. The sample estimate of vacant housing units is controlled to an independently derived estimate for four type-of-vacant cells for each region. All third-stage factors are calculated in a similar manner using the following ratio:

$$\frac{\text{Independently derived estimate of housing units in a cell}}{\text{AHS sample estimate of housing units in that cell}}$$

For occupied units, we derive the numerator of a factor in three steps. First, the Population Division computes an independent estimate of total housing units based on 1990 adjusted census data. Then, we determine the occupied portion of this independent control based on the Current Population Survey distribution for the third-stage occupied cells.

For vacant units, we allocate the vacant portion of the independent control to the distribution of vacant units from the Housing Vacancy Survey (HVS), a monthly vacancy survey conducted by the Bureau of the Census as part of the Current Population Survey.

The denominator of a factor is obtained by summing the weights, with all previous factors applied, for all records in a cell. For the Hispanic/non-Hispanic and vacant cells, this is the weight after the second stage of the ratio estimation procedure. For the Black/non-Black cells, this is the weight after the Hispanic/non-Hispanic portion of the third stage of the ratio estimation procedure. The third-stage ratio adjustment procedures based on the 1980 census data and the 1990 census data are similar. The third-stage ratio estimation factors for the 1989 AHS-National based on the 1980 census data are given in table 7.8.

Raking Procedure

The second and third stages of the ratio estimation procedure are iterated to bring the AHS sample estimates into closer agreement with all independent estimates used.

Table 7.8. **Third-Stage Ratio Adjustment Factors: 1989 AHS-National**

	Northeast	Midwest	South	West
Owner				
Non-Hispanic	1.0050	0.9792	0.9771	0.96307
Hispanic	1.0233	0.9677	0.9603	1.0396
Renter				
Non-Hispanic	0.9951	1.0238	1.0444	1.0224
Hispanic	0.9639	1.7624	1.0955	1.1474
Owner				
Non-Black	0.9991	0.9946	0.9909	1.0056
Black	1.0387	1.1195	1.0759	0.9185
Renter				
Non-Black	0.9440	0.9315	0.9408	0.9514
Black	1.0948	1.1309	1.0624	1.0854

Source: Waite (1990a).

The numerators of the factors are the same ones used previously. The denominators of the factors in this iterative process are obtained by summing the existing weights on all records in a cell. For example, for the second stage of the ratio estimation procedure, the existing weight after the third stage of the ratio estimation procedure from the previous iteration is used. The final weight that results from all iterations is used to produce the tabulations for the AHS report. Further details of the raking procedure used are given in Waite (1990a).

ESTIMATION FOR AHS-MS

The AHS-MS estimates of the characteristics of the housing inventory are produced using a multistage ratio estimation procedure. The basic weight for each interviewed sample housing unit represents the correct probability of selection for each unit. This basic weight = base weight (that is, reciprocal of the probability of selection) x duplication control factor. Before the implementation of the ratio estimation procedure, the basic weight for each housing unit was adjusted to account for Type M and Type A noninterviews.

Type M Noninterview Adjustment

The Type M noninterviews are sample units that were dropped because of selection by another survey or because of permit unavailability. These noninterviews occur in (a) the 1980-based permit-issuing area universe, (b) the 1980-based nonpermit-issuing area universe, and (c) the 1980-based new construction universe.

The adjustment was done separately for the above universes for the central city and balance for each metropolitan area. The adjustment was equal to the following:

$$\frac{\text{AHS-MS sample estimate of 1980 housing units in the cell} + \text{Weighted count of Type M noninterviewed housing units}}{\text{AHS-MS sample estimate of 1980 housing unit in the cell}}$$

Type A Noninterview Adjustment

Type A noninterviews are sample units for which (a) occupants were not home, (b) occupants refused to be interviewed, or (c) occupants were unavailable for some other reason.

The adjustment was done on occupied units and was computed separately for (a) units in the 1980-based permit-issuing area universe, (b) new construction, and (c) all other housing units (this includes the 1970-based permit-issuing universe, the 1970-based and 1980-based nonpermit-issuing universes and the 1970-based new construction housing units built prior to the last survey).

For units in the 1980-based permit-issuing universe a Type A noninterview adjustment factor was computed separately, for each of the 62 strata used in the sample selection process, by 1980 central city and balance. For new construction units, a Type A noninterview adjustment factor was computed separately, by central city and balance. For all other units, a Type A noninterview adjustment factor was calculated separately by tenure and 1970 central city and balance for each of the following:

1. Twenty-four noninterview cells for sample housing units from the permit-issuing universe. Each cell was derived from one or more of the 50 different strata used in the 1970-based permit-issuing universe for selecting the sample.
2. One noninterview cell for new construction housing units.
3. One noninterview cell for mobile homes or trailers from the permit-issuing universe.
4. One noninterview cell for units that were not mobile homes or trailers from the nonpermit-issuing universe.
5. Three noninterview cell for units from the coverage improvement universe.
6. One noninterview cell for units classified as vacants at the time of the 1970 census.
7. One noninterview cell for units classified as group quarters at the time of the 1970 census.

Within a given cell, the Type A noninterview adjustment factor was equal to the following ratio, using the basic weight times the Type M noninterview adjustment factor for the sample weight:

$$\frac{\text{Weighted count of interviewed housing units} + \text{Weighted count of Type A noninterviewed housing units}}{\text{Weighted count of interviewed housing units}}$$

Ratio Estimation Procedure for the 1970-Based Permit-Issuing Universe

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each 1970-based permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

$$\frac{\text{1970 census count of housing units from the 1970-based permit-issuing universe in the corresponding cell}}{\text{AHS-MS sample estimate of the 1970-based housing units from the 1980-based permit-issuing universe in the corresponding cell}}$$

For each metropolitan area, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file (long forms) of housing units enumerated in areas under the jurisdiction of permit-issuing offices.

The denominators of the ratio estimation factors were then obtained from weighted estimates of all the AHS-MS sample housing units from the 1970-based permit-issuing universe, using the existing weight (that is, the basic weight times the Type A noninterview adjustment). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation cells. This ratio estimation procedure was introduced to correct the probabilities of selection for samples, in each of the strata used in the sample selection of the 1970-based permit-issuing universe. Prior to the AHS-MS sample selection within each metropolitan area, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS-MS sample. Since the number of housing units deleted from the AHS-MS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata was introduced during the sample selection process.

Ratio Estimation Procedure for the 1980-Based Permit-Issuing Universe

The following ratio estimation procedure was employed for all sample housing units from the 1980-based permit-issuing universe. This factor was computed separately for all sample housing units within each 1980-based permit-issuing universe noninterview cells mentioned previously. The ratio estimation factor for each cell was equal to the following:

$$\frac{\text{1980 census count of housing units from the 1980-based permit-issuing universe in the corresponding cell}}{\text{AHS-MS sample estimate of the 1980-based housing units from the permit-issuing universe in the corresponding cell}}$$

For each metropolitan area, the numerator of the ratio was obtained from the 1980 Census of Population and Housing 100-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominator of the ratio was obtained from weighted estimates of all the AHS-MS sample housing units within the corresponding ratio estimation categories using the existing weight (that is, the basic weight times the Type M noninterview adjustment factor times the Type A noninterview adjustment factor).

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation categories.

The ratio estimation procedure was introduced to adjust the sample estimate in each of the strata used in the sample selection of the 1980-based permit-issuing universe to an independent estimate (1980 census count) for the strata. This adjustment was necessary since after the sample selection procedure (possibly during materials preparation) some units had to be dropped from sample (for example, as a result of giving up a unit to another survey).

Ratio Estimation Procedures

For the three ratio estimation procedures described below, each metropolitan area was subdivided into geographic areas consisting of a combination of counties.

Mobile home ratio estimation.

$$\frac{\text{Independent estimate of mobile homes for the corresponding geographic subdivision of the metropolitan area}}{\text{Sample estimate of mobile homes for the corresponding geographic subdivision of the metropolitan area}}$$

The numerator of this ratio was determined using census data. The denominator was obtained using the existing weight of AHS sample mobile home units.

Independent total housing unit ratio estimation without mobile homes. This ratio estimation procedure was used in conjunction with the Mobile Home Ratio Estimation procedure.

$$\frac{\text{Independent estimate of total housing inventory (excluding mobile homes) for the corresponding geographic subdivision of the metropolitan area}}{\text{Sample estimate of the total housing inventory (excluding mobile homes) for the corresponding geographic subdivision of the metropolitan area}}$$

The numerator of this ratio was determined using census data. The denominator was obtained using the existing weight of AHS sample units (excluding mobile homes).

Independent total housing unit ratio estimation with mobile homes.

$$\frac{\text{Independent estimate of occupied housing inventory for the corresponding geographic subdivision of the metropolitan area}}{\text{Sample estimate of the occupied housing inventory for the corresponding geographic subdivision of the metropolitan area}}$$

The numerator of this ratio was determined using census data. The denominator was obtained by using the existing weight of AHS sample units.

The computed ratio estimation factors were then applied to all appropriate housing units in the corresponding geographic area of each metropolitan area, and the resulting product was used as the final weight for tabulation purposes.

The decision regarding which of the above mentioned ratio estimation procedures to use was based on the availability of reliable independent controls as well as the size of the mobile home inventory within a metropolitan area. In addition, the decision was based on the magnitude of the mobile home ratio estimation factor within a given metropolitan area.

Note that in even years, the AHS-MS estimates are based on AHS-MS only. But in odd years the AHS-National sample in each AHS-MS area in that year is combined with the AHS-MS sample in the area to produce published AHS-MS estimates. The specification for combined sample weighting for odd years is given in Waite (1990e).

QUALITY CONTROL OF THE ESTIMATION PROCEDURE

At each step in the estimation procedure an extensive verification is built-in to ensure that the results are reasonable and consistent with the requirements of that step. This verification operation includes the production and thorough review of the appropriate output from each process to ensure that the process is being implemented correctly. Any discrepancies identified in this review are then corrected before the next step is implemented.

IMPACT OF ESTIMATION ON DATA QUALITY

For both AHS-National and AHS-MS, the effect of this ratio estimation procedure as well as the overall estimation procedure was to reduce the sampling error for most statistics below what would have been obtained by simply weighting the results of the sample by the inverse of the

probability of selection. Since the housing population of the sample differed somewhat, by chance, from the national or metropolitan area as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, is brought into agreement with known good estimates of the national or metropolitan area housing population.

The first-stage ratio adjustment in the AHS-National reduces the contribution to the variance arising from the sampling of NSR PSU's. The same first-stage factor is used for each survey year until a new sample is selected (usually every 20 years). The use of the same first-stage factors is less efficient at the end of the decade than at the beginning, but there is evidence that the variance is not increased by this practice. The second-stage factors in the AHS-National indicate the degree to which AHS sample estimates are adjusted based on the independent estimates. Bias existing in independent estimates will result in bias in AHS estimates. As with the second-stage factors, the accuracy of the resulting estimates after third-stage ratio adjustments is dependent upon CPS's adjustments for undercoverage and nonresponse.

There is no known unbiased method of adjustment for nonresponse and undercoverage. When adjusting for noninterviews, it is assumed that responses from noninterviewed HU's with the same key characteristics as interviewed HU's would be similar. However, biases exist in the estimates to the extent that responses from noninterviewed HU's or persons have different characteristics than interviewed HU's or persons.

HISTORICAL COMPARISONS

Each home in the AHS sample represents a large number of other homes. The numbers are adjusted so that the total in the survey matches independent estimates of the total number of homes. For 1991, these independent estimates are based on the 1990 Census of Housing, plus changes since then. The 1990-based weighting produces on average, numbers that are about 2.5 percent lower than 1980-based weighting. This effect is not equally distributed among all types of units. The table 7.9 shows the effects of the weighting change by region.

Table 7.10 presents counts of occupied homes using 1990-based weighting. This weighting is consistent with the weighting used to produce the 1991 detailed tables in chapters 1 through 10 of the Current Housing Report H150/91. These data should be used when measuring the change in the size of the occupied inventory over time. These data provide the most accurate count of the total number of occupied homes for the years 1985, 1987, and 1989.

The appendix C of the H150/91 report provides a description of historical changes that have occurred in the American Housing Survey since its beginning in 1973. It also provides appropriate tables that should be used when making comparisons over time for specific characteristics.

Table 7.9. **Difference Between 1980- and 1990-Based Weighting as a Percent of 1980-Based**

Type of unit	United States	North-east	Mid-west	South	West
Total housing units	**2.5**	**3.6**	**2.7**	**2.0**	**1.8**
Occupied	2.4	3.5	2.7	2.0	1.7
Built 1980 or later	0.1	0.0	0.1	0.1	0.1
Built before 1980	2.9	3.9	3.1	2.6	2.2
Vacant	2.9	4.6	2.8	2.4	2.4

Table 7.10. **Occupied Housing Units Using 1990-Based Weighting: 1985, 1987 and 1989**

(Numbers in thousands)

Characteristic	1985		1987		1989	
	Owner	Renter	Owner	Renter	Owner	Renter
United States	**54,394**	**31,279**	**56,649**	**31,885**	**58,193**	**32,809**
Northeast	10,922	7,106	11,418	7,089	11,660	7,011
Midwest	14,226	7,242	14,696	7,133	15,122	7,234
South	19,217	9,876	19,985	10,190	20,627	10,694
West	10,030	7,056	10,550	7,472	10,784	7,870
Race						
White and other	50,222	25,866	52,323	26,253	53,772	26,924
Black	4,172	5,413	4,326	5,632	4,420	5,885

Chapter 8.
Sampling Errors

INTRODUCTION

Estimates derived from AHS data are subject to sampling error because only a portion of the population, the sample, is observed. Sample data can be used to estimate the sampling variance of any sample estimate. Because of cost, the Census Bureau estimates sampling errors for selected items and uses these estimates to develop the values of parameters for use in generalized variance estimates (GVE's). The GVE's can be used by the Census Bureau and by other users of AHS data to estimate the sampling variances associated with any estimate. The standard error of an estimate, as calculated by the Census Bureau, measures not only the sampling error associated with the sampling plan used in AHS, but also partially measures the effect of some nonsampling errors in response and enumeration. It does not measure any systematic biases in the data.

This chapter is not intended to provide detailed information on how to compute sampling errors and construct confidence intervals for specific items. For that purpose, users should refer to the source and accuracy statements that appear in AHS publications (series H150 and H170).

ESTIMATION OF SAMPLING ERRORS

Several methods, including balanced repeated replications (BRR), Taylor series linearization (TSL), jackknife repeated replications (JRR), and random groups (RG) (see Cochran, 1977, Wolter, 1985) have been developed over the years to compute sampling variances from complex surveys like AHS . Robert Fay (1984, 1989) of the Census Bureau has developed a modification of the replication method to improve the stability of the variance estimator. This method, like BRR and other resampling methods, permits the computation of design-based estimates of variance using one simple formula for all kinds of statistics, both simple and nonlinear and other analytically complex statistics. The TSL method is generally computationally efficient, but requires derivation of an appropriate variance formula for each statistic. The Census Bureau used Fay's method for estimating variances directly for selected items for the 1985 AHS-National. A collapsed stratum variance estimator was used for NSR strata by pairing sample PSU's with similar stratum characteristics. Segments from one PSU were assigned to one half-sample and the segments from the other PSU were assigned to the other half-sample. Self-representing (SR) PSU segments were divided into pseudo-PSU's for variance estimation. Within each pseudo-PSU segments were assigned to two half-samples. The variances were computed using 48 half-sample replicates.

The estimates of variances thus computed are biased for two reasons:

1. The AHS design with one sample PSU per stratum precludes an unbiased estimation of variance. Strata were collapsed to develop half-sample replicates. This "collapsed stratum" procedure overestimates the between-PSU variance component.

2. To simplify the variance computation procedures, only 48 half-sample replicates rather than a balanced set of half-sample replicates were used. Further, data for the 48 half-sample replicates were not reweighted and, as a result, these replicates did not reflect the full benefit of the second-stage and third-stage ratio adjustment procedures, in which estimates are adjusted to population totals.

GENERALIZED VARIANCE ESTIMATES (GVE'S)

The variances estimated directly for a selected set of items are used to generalize variances. Generalized variance estimates (GVE's) are used because

1. It would be impractical to compute and/or publish sampling errors for every estimate.

2. The generalized variances give some stability to the estimates of error.

The following equation is used in generalizing the variances:

$$V_x^2 = a + b/x$$

where V_x^2 is the relative variance, the square of the coefficient of variation (standard error/estimate), of the estimate x, and a and b are two parameters, fitted by the least squares method to a set of observed estimates and their computed relative variances. To develop the a and b used in obtaining the generalized standard error tables, a set of estimates of housing characteristics covering a wide numerical range is selected. Through an iterative process,

the estimates and their corresponding relative variances are used to estimate a and b. With the derived a and b, a generalized standard error table for estimates of level is developed.

The Census Bureau uses the GVE's and the estimated parameters a and b in two ways. When analytical statements based on AHS estimates are published, all actual or implied comparisons are tested for statistical significance. For example, a statement that two estimates are different will not be made unless their estimated difference is at least 1.645 times its standard error as determined by using the GVE's. Such a result means that the difference is statistically significant at the 10-percent level.

The Census Bureau also includes GVE's in various publications along with explanations of how to use them. The Appendix B, ''Errors and Source of the Estimates'' in the Current Housing Reports (H150 for AHS-National) provides standard error formulas and illustrates how to compute standard errors for estimates of levels, percentages, ratios, differences between two estimates, and medians. With this information, users may easily calculate estimates of standard errors for any statistics computed from public use files or obtained from published reports. Similar information is also provided in publications in the H151 series for AHS-National Supplements, and the H170 and H171 AHS-MS reports. An example selected portions of the source and accuracy statement from the Current Housing Reports (H150/91) is provided in exhibit 8.1 on the following page.

ESTIMATION OF SAMPLING ERRORS FOR AHS-MS

The Census Bureau used the ultimate cluster method (see Hansen, Hurwitz and Madow, 1953) for estimating variances directly for selected items beginning with the 1984 AHS-MS and the random group method prior to 1984. In the ultimate cluster method, the sample cases associated with each of the sample hits were assigned to unique clusters within four types of sampling universes. These sampling universes were (1) the 1970-based permit-issuing universe which had an expected cluster size of two for each hit, (2) the 1980-based permit-issuing universe which had an expected cluster size of one for each hit, (3) the 1980-based new construction universe which had an expected cluster size of two for each hit, and (4) the 1970-based new construction, 1970-based nonpermit and 1980-based nonpermit universes which have an expected cluster size of four for each hit. Squared deviations among the cluster totals were then computed within each of these four different types of sampling universes and then summed over the four types. In the random group method (see Wolter 1985) the sample was randomly divided into 49 groups and squared deviations among these random group totals were computed.

Exhibit 8.1. **Example of a Source and Accuracy Statement from the Current Housing Reports (H150/91) (only selected portions are shown)**

Appendix B. Errors and Source of the Estimates

SAMPLING AND NONSAMPLING ERRORS

The accuracy of the estimates contained in this report depends on (a) the sampling and on sampling error, as measured by the error formulas in tables 1a through 1c, (b) biases, and (c) other nonsampling errors not measured by the error formulas.

Below is an explanation of sampling and nonsampling error associated with the American Housing Survey (AHS).

Sampling Errors

Sampling error reflects how estimates from a sample vary from the actual value. (Note: By the term ''actual value,'' we mean the value we would have gotten had all housing units been interviewed, under the same conditions, rather than only a sample.)

Suppose based on responses from the sample households we estimate there to be 1,300,000 housing units with a certain characteristic. Because we only interviewed a sample of all households there is a certain amount of ''sampling error'' in this estimate. Because of the sampling error, if we conclude the actual value is between 1,263,000 and 1,337,000 (a 50-percent confidence interval), there is only a 50 percent chance we'll be correct.

The formulas in tables 1a through 1c allow you to compute a range of error such that there is a known probability of being correct if you say the actual value is within the range. The error formulas are approximations to the errors. They indicate the order of magnitude of the errors rather than the actual errors for any specific characteristic. To construct the range, add and subtract the error computed from the formulas to the publication estimate.

The letter ''A'' in the formula represents the publication estimate. Use the number as it appears in the publication (that is, do not multiply it by 1,000).

The letter ''Z'' determines the probability that the actual value is within the range you compute. The larger the value of Z, the larger the range, and the higher the odds the actual value will be in the range. The following values of Z are most commonly used:

Value of Z	Meaning
1.00	There is a 67-percent chance you will be correct if you say the actual value is in the range you compute.
1.60	There is a 90-percent chance you will be correct if you say the actual value is in the range you compute.
1.96	There is a 95-percent chance you will be correct if you say the actual value is in the range you compute.
2.58	There is a 99-percent chance you will be correct if you say the actual value is in the range you compute.

Note that if Z = 1.00, the formula computes the standard error. Ranges of 90 and 95 percent are commonly used. The range of error is also referred to as the confidence interval since there is a certain level of confidence that the actual value is within the interval.

The numbers in this book are printed in thousands (that is, 21 printed in the book means 21,000 homes). The errors are also computed in thousands (that is, do not multiply the number in the publication by 1,000 before computing the error).

For example, the book shows 1,300 elderly households of a certain type (meaning 1,300,000 households since the publication number is in thousands). To compute a 90-percent confidence interval, you would use the first formula in table 1a, and you would compute the error as follows:

$$Z \times \sqrt{(2.288 \times A) - (.000022 \times A^2)}$$

$$1.60 \times \sqrt{(2.288 \times 1{,}300) - (.000022 \times 1{,}300^2)}$$

$$1.60 \times \sqrt{2{,}977.4 - 37.18} = 87$$

There is a 90-percent chance you will be correct if you conclude the actual value is 1,300 plus or minus 87, or in the range 1,213 to 1,387 (which means 1,213,000 to 1,387,000 since the numbers are in thousands).

If the estimate involves two characteristics from tables 1a through 1c, use the formula with the larger first number under the square root. For example, for mobile homes in the South, use the formula for the South since 2.435 is larger than 2.076.

Completeness Rates

Table 3 shows the completeness rates for items from chapter 2 in the publication. The rates indicate what percent of the publication estimates are based on actual responses. The rates for the individual categories of items (for example, income) take the following sources of incomplete data into account:

Item nonresponse (that is, imputation)

Household nonresponse (for example, refusals)

Incomplete coverage (see second and third stage of ratio estimation)

The rates in table 3 are sorted from the lowest rate to the highest for total occupied units.

[redacted]

Chapter 9.
Comparison of AHS With Other Data

INTRODUCTION

We can compare some AHS items to data from other sources, like the census and other surveys, to assess nonsampling errors in AHS. In chapter 5, AHS data have been compared with census data to find differences in year built, units in structure, and tenure items. In this chapter, we provide comparisons of AHS utility costs with data from the Residential Energy Consumption Survey (RECS) and income data with independent estimates, and income and poverty data with Current Population Survey (CPS) estimates.

COMPARISON OF AHS UTILITY COSTS WITH RECS

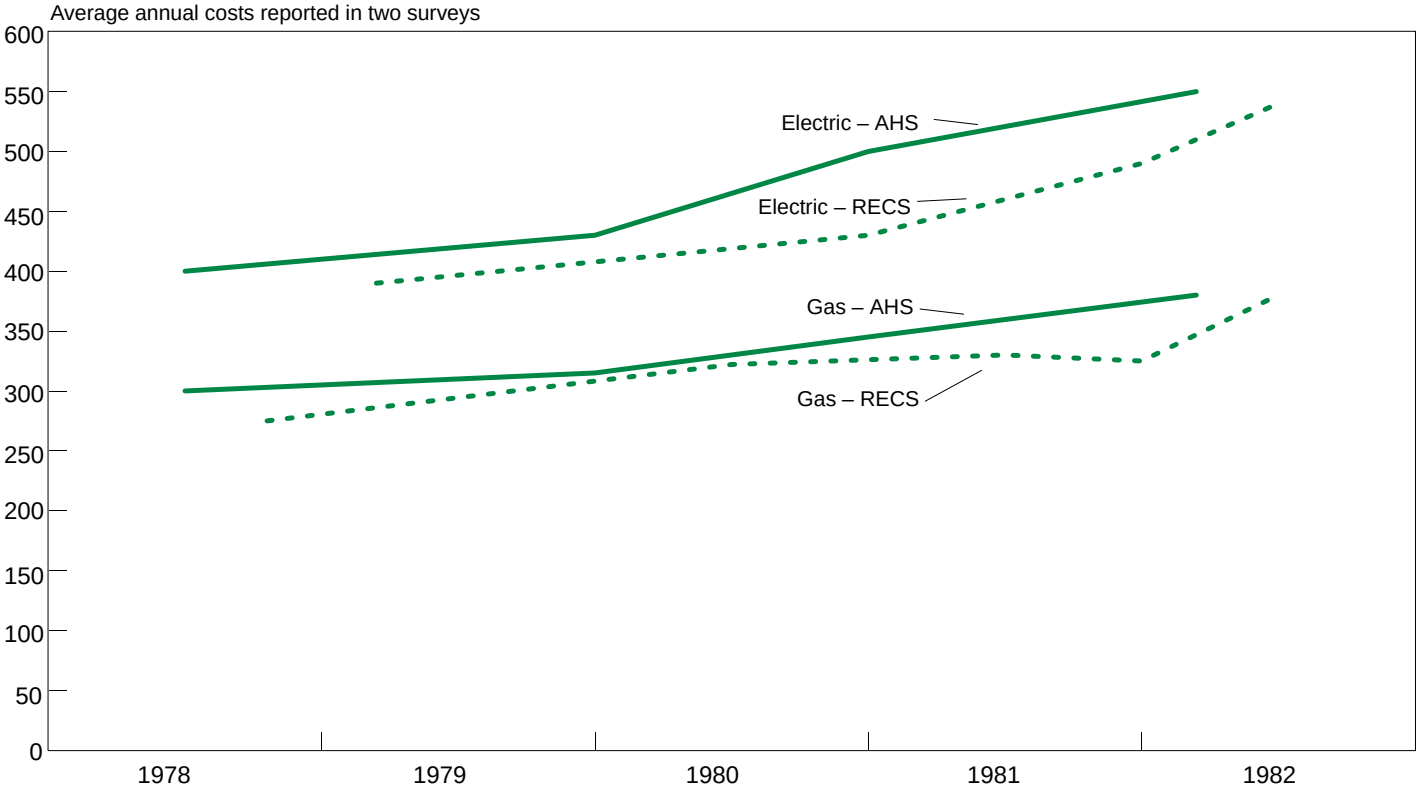
The Residential Energy Consumption Data Survey (RECS) conducted by the Department of Energy collects utility costs data from utility company records. RECS data are, therefore, more accurate than AHS data provided by household respondents. A comparison of AHS utility costs with RECS data is provided in figure 9.1.

This figure clearly shows that AHS reports higher utility costs than the Residential Energy Consumption Survey. The discrepancy is fairly consistent over time, and data not presented here show it is also consistent for single-family detached homes. A plausible reason for the higher AHS figures is that households are more concerned about and, therefore, over-emphasize high-cost months when they mentally average their bills for the AHS field representative.

Figure 9.1

A Comparison of AHS and RECS Utility Costs

Average annual costs reported in two surveys



Source: Energy Information Administration, *Consumption Expenditures, April 1981 through March 1982, Part 1; National Data,* Washington, Government Printing Office, 1983 (and earlier editions), and HUD special tabulations.

Note: This comparison applies to AHS procedures prior to 1989 when procedures were changed to improve AHS utility cost data.

EVALUATION AND RESEARCH ON UTILITY COSTS

Previous evaluations of decennial census data have also indicated that the estimates of the average monthly cost of gas and electricity are subject to relatively large response biases (net overreporting) and that the size of the bias varies considerably from area to area. A utility cost study was conducted during the 1980 census in eight cities to evaluate potential gain in accuracy of responses if respondents are provided with an average monthly bill for 12 months prior to the census. Tippett and Takei (1983) provide the results of this method test involving six cities where half of the sample households were provided by utility companies with average monthly utility cost for the 12 months prior to the 1980 census along with March 1980 bills. The actual costs incurred, as reported by the utility companies, were then compared with the amounts reported by the same households on the 1980 census long-form questionnaire.

Overall, census respondents tended to overreport their cost of gas more than they overreported their cost of electricity. Also renter-occupied households tended to overreport their cost of gas and electricity more than owner-occupied households. For electricity, the improvement resulting from the notification was 22.6 percent for renters and 38 percent for owners, but the notified census respondents still overreported their cost by 15.2 percent for owners and 26.0 percent for renters. For gas, the improvement was 26.7 percent for renters and 48.4 percent for owners, but the notified census respondents still overreported their cost by 29.7 percent for owners and 41.2 percent for renters.

Mortgaged households did a better job of reporting their cost of electricity than the nonmortgaged households, and there was only a slight difference between mortgaged and nonmortgaged households when reporting the cost of gas. For electricity the improvement resulting from the notification was 46.9 percent for mortgaged households and 20.8 percent for nonmortgaged households, but the notified census respondents still overreported their cost of electricity by 11.9 percent for mortgagers and 22.9 percent for nonmortgagers. For gas, the improvement was 46.9 percent for mortgagers and 52.7 percent for nonmortgagers, but the overreporting was 29.6 percent for mortgagers and 28.5 percent for nonmortgagers.

Providing customers with their average monthly cost of electricity did make a significant improvement in estimating the shelter costs of owner-occupied and mortgaged units, but there was no improvement apparent in estimating the shelter cost of homeowners with no mortgage and only some evidence that the improvement was significant for estimating the gross rent for renters. However, providing customers with their average monthly cost of gas did make a significant improvement for both the shelter cost of homeowners, and gross rent of renters.

Another problem with utility costs in AHS is that sometimes respondents provide a combined cost for two or more utilities. Williams (1981) describes problems with combined utility cost in the Memphis SMSA, where many residents receive a monthly utility bill on which charges for several utilities are combined. Although the charges are itemized, respondents frequently report only the total amounts when answering the AHS items on utility cost. Therefore, the combined amount for anywhere from two to five different utilities may be entered on the line for monthly cost of electricity.

In another study, Williams (1982) examines the utility costs reported in Panel 6 for the Albany, Madison, and Spokane SMSA's. These SMSA's exhibited relatively high electricity costs and low gas costs, one indication that a large number of units reported combined utility costs. Three characteristics which had the most impact on fuel costs were all related to a unit's location. In order of importance they are: region, PSU, and SMSA/non-SMSA. Utility companies' billing policies can vary widely in adjacent cities (and by extension adjacent PSU's). Williams recommends that for AHS-MS data should be grouped at least by inside or outside central city when allocating combined utility costs in order to more accurately reflect the operating factors for combined utility reporting.

Bateman and Williams (1982) recommended using the ''hot-deck'' allocation procedure for choosing a donor household as is used in allocating such items as income for AHS-National. The ratio of individual utility costs for the donor unit would be applied to the combined amount reported in the other households.

In 1989, an attempt was made to collect actual utility cost data from utility companies in two metropolitan areas (Washington, DC and Minneapolis-St. Paul) but it failed to obtain cooperation from most of the area utility companies.

The estimation of utility costs for AHS-National by regression using monthly utility cost data from the RECS public use file and some common RECS/AHS housing characteristics as independent variables was researched by Sliwa (1988a, 1988b). Sliwa (1989) provided specifications for deriving annual costs for electricity and natural gas.

Beginning with the 1989 AHS-National, the Census Bureau changed the procedures for collecting and processing utility cost data. In an attempt to improve the utility cost data, respondents were asked to consult their records to provide cost data for four specific months in the past year. This request was included in the letter (exhibit 4.1, chapter 4) mailed to each sample address before enumeration. Electricity and natural gas costs for the months of January, April, July, and December were asked. If the respondents could not provide the costs for at least two of the specified months, the respondents were asked to estimate the average monthly cost for the utility for the last 12 months.

During processing, if the utility costs for at least two (one for recent movers) of the specified months were received, the regression approach developed by Sliwa (1989) was used to estimate the average monthly cost of the utility. If the respondent did not report costs for enough of the

specified months, the average monthly estimate provided by the respondent was used. If the respondent did not provide a monthly estimate, a ''hot-deck'' allocation was used to obtain a monthly estimate. The ''hot-deck'' did not use any of the regression estimates to fill the matrix. After editing the utility cost data for all records, the average monthly cost data for the records that did not use the regression method were adjusted by:

1. Multiplying the sum of the number of cases using the regression method plus those that did not use the regression method by the average monthly cost based on the most recent RECS data, adjusted for inflation.

2. Subtracting the sum of the monthly costs for cases using the regression method from the result of one above.

3. Dividing the result of two above by the sum of the monthly costs for cases that did not use the regression method.

4. Multiplying the results from three above to the average monthly cost value for each case that did not use the regression method.

This is done separately for electricity and natural gas.

COMPARISON OF AHS INCOME WITH INDEPENDENT ESTIMATES

It is well known that income statistics derived from household surveys are generally biased due to response errors. Respondents tend to underestimate income. We can assess the accuracy of AHS income data by comparing with independent estimates. Such a comparison as reported in HUD and Bureau of the Census (1990) follows.

Independent estimates of income from GNP accounts, the Social Security Administration, the Veterans Administration, etc., and CPS and AHS income estimates are shown in table 9.1.

AHS figures are lower than the independent estimates for total income and for every category other than self-employment income. The Current Population Survey (CPS) is done by the Census Bureau for the Bureau of Labor Statistics. It is also low but comes closer to the independent estimates. This may be largely due to the differences in income questionnaires and timings of CPS and AHS. More detailed and extensive questions about income sources and amount by source are asked in CPS than in AHS. Also, the CPS March supplement for income coincides with income tax time when respondents are more aware of nonwage incomes like interest, dividends, etc.

Table 9.1. **Money Income of All U.S. Households**
(Billions of dollars)

	Independent estimate (dollars)	CPS (dollars)	AHS (dollars)	AHS as percent of independent estimate
Total money income ...	**[1]2,402**	**2,201**	**2,073**	**86**
Wages or salaries	1,632	1,161	1,505	92
Interest	221	99	67	30
Social Security, railroad retirement	155	142	139	90
Nonfarm self-employment ...	104	120	142	137
Dividends	60	27	[2]38	63
Estates and trusts	(NA)	7	(NA)	(NA)
Federal and military retirement	35	32	33	94
State and local government retirement	21	13	(NA)	(NA)
Private pensions and annuities	55	35	27	49
Net rent and royalties	34	17	[2]23	68
Unemployment compensation.....	26	20	18	69
AFDC	14	11	(NA)	(NA)
SSI	9	8	17	189
Other public assistance	(NA)	2	(NA)	(NA)
Workers' compensation	14	7	5	36
Veterans' payments	14	9	[2]13	93
Farm self-employment	9	10	25	278
Alimony and child support ...	(NA)	8	8	(NA)
Regular contribution from people	(NA)	5	5	(NA)
Other money income	(NA)	14	9	(NA)

(NA) Not available.

[1]Excludes 5 categories, shown as (NA). There are other differences such as the exclusion of children's income ages (0-14) from CPS and AHS, military households from CPS, and group quarters from AHS.

[2]AHS comes closer to independent estimate than CPS does. This is considered desirable, but even the independent estimates contain unknown amounts of errors.

Reference period months ending 12/83, 12/83, 10/83 for independent CPS and AHS estimates.

Source: Census Series P-60, No. 151, p. 170 and HUD special tabulation. (Since the AHS public use tape did not distinguish among amounts of $50,000 or more, they have each been treated as $60,000.)

COMPARISON OF AHS AND CPS INCOME REPORTING

Recently, Williams (1992) provided an extensive comparison of the data on income that were collected in the 1989 AHS-National and the March 1990 CPS. This comparison addresses the topics:

- The percent of households reporting income by each income source
- The amount of income reported by income source
- The total amount of income reported

This analysis at least partially supports the hypothesis that the AHS income estimates are lower than CPS largely due to the less detailed AHS income questions. Major results of the study follow:

1. The AHS and CPS differences (for households with NO nonrelatives) were concentrated in the nonwage component of income. For the wage and salary portion of income, AHS response rate and amount reported were a little higher than in CPS. For households containing nonrelated persons (where a comparison of the wage and salary and nonwage portions of income was not relevant), the ''undercount'' was most noticeable among the nonrelated persons rather than the householder and relatives. To better match the levels of CPS income, AHS needs to determine how to get respondents to report their nonwage income—particularly, interest and dividends and all other sources—and/or income for persons not related to the householder. These areas are, perhaps not coincidentally, sections of the AHS interview gathered with considerably less detail than the CPS interview.

2. Within the category ''all other sources,'' table 9.2 reveals that educational assistance benefits was one of the more frequently mentioned CPS ''other'' income sources but it is one not specifically mentioned in the AHS interview. Also, disability benefits, though reported less frequently, had the highest value of the ''other'' income sources. It is also not mentioned in AHS.

   Eighty-eight percent of the CPS households with ''all other income'' obtained all the income from a single source. Another 11 percent received money from just two of these sources. Very few households received income from more than one of these ''all other'' sources. Therefore, if the particular component of nonwage ''all other sources'' is not specifically mentioned in the questionnaire, the household is likely to record no income from ''all other'' nonwage sources.

   Adding suitable phrases to the AHS ''all other income'' question (or inserting a new question) in order to pick up educational and disability benefits may help close the gap between AHS and CPS income levels. Likewise some of the underreporting in the ''interest and dividends'' category may be reduced by expanding the question wording to include examples of the sources from which the households would expect to receive interest and dividends.

3. The problem with AHS underreporting nonrelatives' income seems somewhat different. Since 1985, the AHS interview has placed the sole question on nonrelative income at the very end of the interview in order to make it easier to get self-reporting (and hopefully better estimates) of nonrelatives' income. It was not possible to determine how much of nonrelative income was self-reported in the 1989 AHS-National. However, since CPS does not attempt to obtain self-reporting for nonrelatives' (as well as relatives') income and tends to get higher amounts for both, this strategy may not be necessary. It may be more productive to revert to the former AHS questionnaire design in which nonrelative income is obtained using the same series of questions as was used for the relatives' income.

Table 9.2. **Components of ''All Other Income'' Category From CPS for Households With NO Nonrelatives**

	Percents of households with this source	Median amount received (households with any income from source) (dollars)
Disability benefits	1.9	5,961
Educational assistance benefits	6.2	1,945
Financial assistance payments	1.3	2,902
Other income payments	2.2	862
Unemployment compensation benefits	5.7	1,624
Veterans' payments	2.8	2,513
Workers' compensation	2.3	2,442

Source: Williams (1992).

COMPARISON OF AHS AND CPS POVERTY DATA

This comparison (Williams, 1995) documents AHS and CPS poverty levels between 1985 and 1993. It concentrates on procedural differences between these surveys and their subsequent impact on poverty level reporting. The Census Bureau has introduced three major changes in AHS data collection or processing related to the poverty data since 1985.

Monthly Moving Poverty Threshold

Beginning in 1989, the AHS has used a set of monthly moving poverty thresholds based on the 12 sets of poverty thresholds for the 12 months prior to the interview. This change was to align the poverty cutoffs more closely with how the income data were collected. The AHS asks for income in terms of the last 12 months as of the date of the interview. This date varies from mid-summer through the end of the year.

The result of this procedural difference has not been measured. In the year it was introduced, the AHS poverty rate did not change from the previous survey period even though the corresponding years of CPS data showed a significant decline. However, the contribution of the new method of applying thresholds to these results is unknown.

Nonwage Income Items

In 1993, the Census Bureau revised the nonwage income section of the AHS questionnaire. The intent was to pick up income sources commonly reported in CPS which

previously had not been specified in the AHS interview. Due, at least in part, to the questionnaire changes, the percent of households with nonwage income rose from 63 percent to 77 percent between 1991 and 1993—a change that should have reduced (assuming positive nonwage incomes) the 1993 AHS poverty rate from what it would have been without the revisions. The actual amount of money the new questions picked up was probably small. One indication of this is that the median nonwage income dropped between 1991 and 1993, from $7,400 to $6,212.

Presence of Lodgers

Also in 1993, a questionnaire change was introduced for the data on the presence of lodgers in the household. The definition of a lodger was expanded to include *all* persons (14 years or older) not related to the householder, who paid rent or part of the housing costs. Due to a consistency edit between presence of lodgers and their payments and householder rental income, the question change also produced an increase in the percent of households reporting rental income. Thus, this revision would tend to reduce the number of AHS households in poverty. However, because of the previously mentioned simultaneous changes in the categories for nonwage income, it is impossible to gauge how much this revision added to a household's income.

Comparison of AHS and CPS Between 1985 and 1993

Table 9.4 outlines the comparable household poverty data from the two surveys since 1985. Although the overall poverty rate offers a good match between AHS and CPS, data users should be cautious in their use of AHS poverty data. The CPS is the Census Bureau's official source of poverty data and the AHS ″poverty″ statistics are approximations of this standard. Furthermore, the CPS data are expressed in terms of families and/or individuals, not households. And, the fit between AHS and CPS data breaks down when comparing at least some subgroups. Users cannot assume that because the overall poverty levels are comparable, that the AHS poverty counts for households in 1-bedroom units (or some other subset of the population) would be the same as CPS would produce.

Table 9.3. **Percent of AHS Households Reporting Nonwage Income by Source: 1991 and 1993**

(In percent. 1993 categories shown in italics)

Source	1991	1993	Difference
Business, farm or ranch	11.8	12.2	0.4
Social security or pensions	30.1	29.8	*–0.3
Interest and dividends	23.3	46.4	23.1
Interest		*44.7*	
Dividends		*17.5*	
Rental income	8.6	12.1	3.5
Welfare, SSI	6.4	6.3	*–0.1
Alimony, child support	4.3	4.6	0.3
Unemployment or workers' compensation, or all other sources	9.6	13.8	4.2
Workers' compensation, disability		*4.0*	
Unemployment, veteran's payments, all other		*10.4*	

*Difference is **not** statistically significant.

Table 9.4. **Households in Poverty, AHS and CPS: 1985 to 1993**

(In thousands)

Year households	Total poverty households		Percent of—	
	AHS	CPS	AHS	CPS
1985	13,266	11,996	15.0	13.6
1987	11,969	11,945	13.2	13.1
1989	12,403	11,369	13.2	12.2
1991 (1980-based)	13,160	12,949	13.8	13.5
1991 (1990-based)	12,836	(NA)	13.8	(NA)
1993	13,787	13,847	14.6	14.2
Difference within surve:y				
1985-87	–1,297	*–51	–1.8	0.5
1987-89	434	–576	*0.0	–0.9
1989-91 (1980-based)	757	1,580	0.6	1.3
1991-93 (1990-based)	951	(NA)	0.8	(NA)
Difference between surveys:				
1985	1,270		1.4	
1987	*24		*0.1	
1989	1,034		1.0	
1991 (1980-based)	*211		*0.3	
1993	*–60		*0.4	

Difference is **not** statistically significant.
(NA) Not available.

Abbreviations

AHS	American Housing Survey
AHS-MS	American Housing Survey-Metropolitan Sample
AHS-National	American Housing Survey-National Sample
AOQL	Average Outgoing Quality Limit
BRR	Balanced Repeated Replications
CATI	Computer Assisted Telephone Interviewing
CMSA	Consolidated Metropolitan Statistical Area
CPS	Current Population Survey
DCF	Duplication Control Factor
DK	Do Not Know
DSD	Demographic Surveys Division
DSMD	Demographic Statistical Methods Division
ED	Enumeration District
FR	Field Representative
GDR	Gross Difference Rate
GQ	Group Quarter
GVE	Generalized Variance Estimate
NHIS	National Health Interview Survey
HHES	Housing and Household Economic Statistics Division
HU	Housing Unit
HUCS	Housing Unit Coverage Study
HUD	Housing and Urban Development
HVS	Housing Vacancy Survey
JRR	Jackknife Repeated Replications
MPP	Moderate Physical Problems
MS	Metropolitan Sample
MSA	Metropolitan Statistical Area
MUS	Multiunit Structure
NC	New Construction
NSR	Non Self-Representing
PMSA	Primary Metropolitan Statistical Area
PSU	Primary Sampling Unit
PV	Personal Visit
QC	Quality Control
RECS	Residential Energy Consumption Survey
RG	Random Groups
RIM	Record/Item Management
RO	Regional Office
SMSA	Standard Metropolitan Statistical Area
SOMHP	Survey of Mobile Home Placements
SR	Self-Representing
TSL	Taylor Series Linearization
UCL	Upper Control Limit
URE	Usual Residence Elsewhere
UTL	Unable-to-Locate

References

Abernathy, J.F. (1987), ''1987 AHS-MS—Results of Research on Wave 1 Preedit Rejects for Units in Structure Inconsistent,'' Internal Census Bureau memorandum to Cannon, November 3.

_____(1991), ''1989 AHS-MS—Results of Research on Regional Office Preedit,'' Internal Census Bureau memorandum to Cannon, July 24.

Bailey, L., Moore, T., and Bailar, B. (1978), ''An Interviewer Variance Study for Eight Impact Cities of the National Crime Survey Cities Sample,'' Journal of the American Statistical Association, 73, 16-23.

Bailar, B., Bailey, L., and Corby, C. (1978), ''A Comparison of Some Adjustment and Weighting Procedures for Survey Data,'' In N. Krishman Namboodiri, ed., Survey Sampling and Measurement, Academic Press, New York.

Bateman, D.V., and Williams, B.T. (1982), ''Implementation of Proposed Combined Utility Cost Imputation for the Annual Housing Survey (AHS),'' Internal Census Bureau memorandum to Young, February 17.

Brooks, C., and Bailar, B. (1978), ''An Error Profile: Employment as Measured by the Current Population Survey,'' Statistical Policy Working Paper 3, Office of Federal Statistical Policy and Standards, U.S. Department of Commerce.

Buckles, E.A. (1981), ''Verification of Reporting of Cooperatives and Condominiums in the 1979 Annual Housing Survey-National Sample,'' Internal Census Bureau memorandum to Koons, May 1.

Bureau of the Census (1968), ''Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: Effects of Interviewers and Crew Leaders,'' Series ER 60, No. 7, U.S. Government Printing Office, Washington, DC.

_____(1978), ''The Current Population Survey: Design and Methodology,'' Technical Paper 40, U.S. Government Printing Office, Washington, DC.

_____(1985a), ''American Housing Survey-National Sample Field Representatives' Manual,'' Internal Census Bureau Manual.

_____(1985b), Evaluating Censuses of Population and Housing, Statistical Training Document, ISP-TR-5, Washington, DC.

_____(1986), ''Content Reinterview Study: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview,'' 1980 Census of Population and Housing, Research and Evaluation Report PHC80-E2, U.S. Government Printing Office, Washington, DC.

Chakrabarty, R.P. (1992a), ''Summary of Nonsampling Errors in the AHS through Reinterview Studies from 1973 through 1985,'' Internal Census Bureau memorandum for the record, July 15.

_____(1992b), ''Study of Building Loss—Vacant Other Data,'' Internal Census Bureau memorandum for the record, December 22.

_____(1993a), ''Number of Interviews and Noninterviews by Type for AHS-MS, 1986-1989,'' Internal Census Bureau memorandum for the record, May 18.

_____(1993b), ''Quality Profile for the American Housing Survey,'' Proceedings of the session on Survey Research Methods, American Statistical Association.

_____(1993c), ''Measurement Errors in the American Housing Survey,'' Bulletin of *International Statistical Institute,* proceedings of the 49th Session. Firenze.

Cochran, W.G. (1977), ''Sampling Techniques,'' 3rd Edition, John Wiley & Sons, New York.

Dinwiddie, J. (1977), ''Results and Analysis of the 1975 Rural Listing Test,'' Bureau of the Census 1975 Rural Listing Test Results memorandum No. 4, November 14.

Fay, R. E. (1984), ''Some Properties of Estimates of Variance Based on Replication Methods,'' Proceedings of the Section on Survey Research Methods, American Statistical Association.

_____(1989), ''Theory and Application of Replicate Weighting for Variance Calculations,'' Proceedings of the Section on Survey Research Methods, American Statistical Association.

Felligi, I.P. (1974), ''An Improved Method of Estimating the Correlated Response Variance,'' Journal of the American Statistical Association, 69, 496-501.

Hansen, M.H., Hurwitz, W.N., and Pritzer, L. (1964), ''The Estimation and Interpretation of Gross Differences and the Simple Response Variance,'' Contributions to Statistics, presented to Professor P.C. Mahalanobis on the occasion of his 70th birthday, Ed. C.R. Rao, Statistical Publishing Society, Calcutta, 111-136.

Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953), ''Sample Survey Methods and Theory,'' Vol 1, John Wiley and Sons, New York.

Harris, J.S. (1984), ''American Housing Survey (AHS)-National Sample Pretest of Area Segment Procedure (Project 4007),'' Internal Census Bureau memorandum to Matchett, December 13.

Hartnett, B. (1985), ''Results of the Verification of Cooperative and Condominium Followup for the Annual Housing Survey-1983 National Sample,'' Internal Census Bureau memorandum to Montfort, March 22.

Hayes, E.A. (1989), ''Electrical Breakdowns—A Test of Alternative Wording in the 1988 AHS-MS Reinterview,'' Internal Census Bureau memorandum for the record, September 18.

HUD and Bureau of the Census (1990), ''Codebook for the American Housing Survey, Database: 1973-1993,'' U.S. Government Printing Office, Washington, DC.

Leadbetter, S.D., Cole, C.G., Meier, A.J., and Huang, H. (1991), ''The Impact of Computer Assisted Telephone Interviewing on Data from the American Housing Survey,'' presented at the Joint Statistical Meeting in Atlanta, GA, August 19.

Loudermilk, C. (1989), ''Data on Discontinued Offices,'' Internal Census Bureau memorandum to Petroni, March 29, with Attachment, ''Permit Office Boundary Problems,'' February 14.

Mahalanobis, P.C. (1946), ''Recent Experiments in Statistical Sampling in the Indian Statistical Institute,'' Journal of the Royal Statistical Society, 109, 325-378.

Masumura, W. (1981), ''Initial Findings of Research on the Validity of Reasons Given for Moving,'' Internal Census Bureau memorandum to Koons, January 13.

Matchett, S.D. (1985), ''Interviewers Instructions for AHS-National Coverage Improvements Segments (Project 0904),'' Internal Census Bureau memorandum for selected AHS interviewers, August 13.

_____(1987), ''Interviewers Instructions for AHS-National Coverage Improvement (Area) Segments (project 0904),'' Internal Census Bureau memorandum for selected interviewers, August 7.

Montfort, E. (1983a), ''Study of Reason-For-Move-To Questions on the 1979 and 1982 National Annual Housing Survey,'' Internal Census Bureau memorandum to Young, January 13.

_____(1983b), ''Coding 1978 Annual Housing Survey National Reason-for-Move Responses,'' Internal Census Bureau memorandum to Young, May 10.

Parmer, R.J. (1986), ''Documentation of the AHS-National Noninterview Research for 1985,'' Internal Census Bureau memorandum for the documentation, April 16.

Parmer, R.J., Huang, H., and Schwanz, D.J. (1989), ''The Impact of Telephone Interviewing on Data from the American Housing Survey,'' Proceedings of the Survey Research Methods Section, American Statistical Association.

Quansey, A. (1986), ''Time Study Conducted During Panel 10 Enumeration of the 1985 AHS-MS (Project 0907),'' Internal Census Bureau memorandum to Cannon, March 10.

Schreiner, I. (1977), ''CPS Reinterview Results from the Listing Check, Check of Noninterview Classifications, and the Household Composition Check for 1976,'' Internal Census Bureau memorandum, April 27.

Schwanz, D.J. (1988a), ''Mobile Home New Construction for 1985 AHS-National,'' Internal Census Bureau memorandum to Montfort, February 3.

_____ (1988b), ''1985 Type-A Unable-to-Locate Rates for the AHS-National Sample Units,'' Internal Census Bureau memorandum for the distribution list, November 14.

Shapiro, G.M. (1980), ''Listing Errors in Area Segments,'' Internal Census Bureau memorandum for the redesign committee, December 8.

Silwa, G.E. (1988a), ''Regression Analysis of Utility Costs Using the 1984 Residential Energy Consumption Survey (RECS) Public Use File,'' Internal Census Bureau memorandum for the distribution list, October 12.

_____(1988b), ''Specialized Regression Analysis of Utility Costs Using the 1984 Residential Energy Consumption Survey (RECS),'' Internal Census Bureau memorandum for the distribution list, November 4.

_____(1989), ''Specifications for Deriving Annual Costs for Electricity and Natural Gas in the American Housing Survey (AHS),'' Internal Census Bureau memorandum to Bartlett, September 29.

Smith, R.T. (1985), ''Analysis of the Reinterview Data from the Annual Housing Survey—SMSA, Group AA-2,'' Internal Census Bureau memorandum for the cc list, August 30.

Statt, R., Vacca, E.A., Wolters, C., and Hernandez, R. (1981), ''Problems Associated with Using Building Permits as a Frame of Post-census Construction: Permit Lag and Identification,'' Proceedings of the Section on Survey Research Methods, American Statistical Association, 226-231.

Taueber, C., Thompson, J., and Young, A.F. (1983), ''1980 Census Data: The Quality of the Data and Some Anomalies,'' Internal Census Bureau memorandum.

Tepping, B., and Boland, K. (1972), ''Response Variance in the Current Population Survey,'' Working Paper 36, Bureau of the Census.

Tippett, J. (1988), ''Tampa AHS-Census Match Study,'' Preliminary Research and Evaluation memorandum, Internal Census Bureau memorandum to Norry, January 13.

Tippett, J. and Takei, R. (1983), ''Evaluation of Reporting of Utility Costs for Selected Cities,'' Preliminary Evaluation Research memorandum No. 59, 1980 Census, August 31.

Waite, P.J. (1990a), ''1989 AHS-National Specifications for Weighting Regular Data,'' Internal Census Bureau memorandum to Courtland, April 26.

_____(1990b), ''Report of the 1985 AHS-MS Reinterview,'' Internal Census Bureau memorandum to Courtland, May 4.

_____ (1990c), ''Results of CPS Reinterview Quality Control for 1988,'' Internal Census Bureau memorandum to Matchett, June 6.

_____(1990d), ''Quality Assurance Results for Keying American Housing Survey—Metropolitan Sample (AHS-MS) May 1989-December 1989,'' Internal Census Bureau memorandum to Matchett, June 6.

_____(1990e), ''AHS-MS: Specifications for Weighting Regular Data from the 1989 MSS,'' Internal Census Bureau memorandum to Courtland, September 11.

_____(1990f), ''Quality Assurance Results for Keying 1989 American Housing Survey-National Sample (AHS-N),'' Internal Census Bureau memorandum to Matchett, September 26.

_____(1993), ''Recommendation on Using Computer Assisted Telephone Interviewing (CATI) in Future Enumerations of the American Housing Survey-National Sample (AHS-N),'' Internal Census Bureau memorandum to Weinberg and Courtland.

Weidman, L. (1988), ''Nonresponse Rates for the American Housing Survey,'' Internal Census Bureau memorandum to Schwanz, April 7.

Wells, G.E. (1982), ''Review of 1981 AHS-National Proxy Interviews,'' Internal Census Bureau memorandum, June 21.

Wetzel, A.J. (1990), ''Quality Assurance of Keying, Derivation of Accept/Reject and Requalification Tables,'' Internal Census Bureau memorandum to Busbery, December 20.

Williams, B.T. (1978), ''The Building Loss—Vacant Other Recheck Program,'' Internal Census Bureau Draft, September 25.

_____(1979), ''The Analysis of 1978 and 1979 Building Loss—Vacant Other Recheck,'' Internal Census Bureau Document.

_____(1981), ''Resolution of Reported Combined Utility Cost, Memphis SMSA, Year cc-1 (1980-81),'' Internal Census Bureau memorandum to Montfort, August 25.

_____(1982), ''Further Research on Combined Utility Cost Reporting in AHS-SMSA,'' Internal Census Bureau memorandum to Montfort, April 6.

_____(1985), ''Results from the Multi-Unit Structures Followup to the 1984 American Housing Survey: Metropolitan Sample (AHS-MS),'' Internal Census Bureau memorandum for the record.

_____(1992), ''Comparison of 1989 AHS and CPS Income Reporting,'' Internal Census Bureau memorandum for the record, May 18.

_____(1995), ''American Housing Survey (AHS) Poverty Data: 1985 to 1993,'' Internal Census Bureau memoranda for the record, October 25.

Wolter, K.M. (1985), ''Introduction to Variance Estimation,'' Springer-Verlag, New York.

Young, A.F. (1982), ''Comparison of 1980 Census Provisional Estimates with AHS Estimates for Selected Characteristics,'' Internal Census Bureau memorandum to Bailar, June 18.