DOE/EIA-0555(94)/1

Energy Consumption Series

Sample Design for the Residential Energy Consumption Survey

August 1994

Energy Information Administration

This publication and other Energy Information Administration (EIA) publications may be purchased from the Superintendent of Documents, U.S. Government Printing Office.

All telephone orders should be directed to:

U.S. Government Printing Office
McPherson Square Bookstore
1510 H Street, N.W.
Washington, DC 20005
(202)653-2050
FAX (202)376-5055
9 a.m. to 4:30 p.m., eastern time, M-F

Superintendent of Documents
U.S. Government Printing Office
Washington, DC 20402
(202)783-3238
FAX (202)512-2233
8 a.m. to 4 p.m., eastern time, M-F

All mail orders should be directed to:

U.S. Government Printing Office
P.O. Box 371954
Pittsburgh, PA 15250-7954

Complimentary subscriptions and single issues are available to certain groups of subscribers, such as public and academic libraries, Federal, State, local and foreign governments, EIA survey respondents, and the media. For further information and for answers to questions on energy statistics, please contact EIA's National Energy Information Center. Address, telephone numbers, and hours are as follows:

National Energy Information Center, EI-231
Energy Information Administration
Forrestal Building, Room 1F-048
Washington, DC 20585
(202)586-8800
Internet E-Mail: INFOCTR@EIA.DOE.GOV
TTY: For people who are deaf or hard
of hearing: (202)586-1181
9 a.m. to 5 p.m., eastern time, M-F

Released for Printing: August 1, 1994

Printed with soy ink on recycled paper

DOE/EIA-0555(94)/1
Distribution Category UC-950

Energy Consumption Series

Sample Design for the Residential Energy Consumption Survey

August 1994

Energy Information Administration
Office of Energy Markets and End Use
U. S. Department of Energy
Washington, DC 20585

This report was prepared by the Energy Information Administration, the independent statistical and analytical agency within the Department of Energy. The information contained herein should not be construed as advocating or reflecting any policy position of the Department of Energy or any other organization.

Contacts

This report was prepared under the general direction of W. Calvin Kilgore, Director of the Office of Energy Markets and End Use (202/586-1617). The project was directed by Lynda T. Carlson, Director of the Energy End Use and Integrated Statistics Division (EEUISD) (202/586-1112) and Nancy L. Leach, Chief of the Residential and Commercial Branch (202/586-1114). Specific information regarding the contents or preparation of this publication can be obtained from the principal author, Robert B Latta (202/586-1385). Organization and production of this report was done by Ivy Harrison and Hattie Ramseur. Questions regarding the Residential Energy Consumption Survey (RECS) may be referred to Wendel Thompson, RECS Survey Manager (202/586-1119). The FAX number for all EEUISD personnel is 202/586-0018.

Contents

Appendices

Tables

Illustrations

Acronym Listing

1.	CBECS	Commercial Buildings Energy Consumption Survey
2.	CPS	Current Population Survey
3.	EIA	Energy Information Administration
4.	EPACT	Energy Policy Act (P.L. 102-486)
5.	HHS	U.S. Department of Health and Human Services
6.	MCD	Minor Civil Division
7.	MOS	Measure of Size
8.	MSA	Metropolitan Statistical Area
9.	NIECS	National Interim Energy Consumption Survey
10.	NSR	Non-self Representing
11.	PPS	Probability-Proportional-to-Size
12.	PSU	Primary Sampling Unit
13.	RECS	Residential Energy Consumption Survey
14.	RSE	Relative Standard Error
15.	RTECS	Residential Transportation Energy Consumption Survey
16.	SSU	Secondary Sampling Unit
17.	TIGER	Topologically Integrated Geographic Encoding and Referencing

1. Introduction

The purpose of this report is to provide detailed information about the multistage area-probability sample design used for the Residential Energy Consumption Survey (RECS). It is intended as a technical report, for use by statisticians, to better understand the theory and procedures followed in the creation of the RECS sample frame. For a more cursory overview of the RECS sample design, refer to the appendix entitled "How the Survey was Conducted," which is included in the statistical reports produced for each RECS survey year. The most recent reports in the RECS series include:

- *Housing Characteristics 1990*, DOE/EIA-0314(90), May 1992 (GPO Stock No. 061-003-00754-6)
- *Household Energy Consumption and Expenditures 1990*, DOE/EIA-0321/1(90), February 1993 (GPO Stock No. 061-003-00795-3).

Overview of the Residential Energy Consumption Survey

The Energy Information Administration (EIA) is mandated by Congress to be the agency that collects, analyzes, and disseminates impartial, comprehensive data about energy; its users, and the purposes for which it is used. To comply with that congressional mandate, EIA collects energy data from the following sectors: residential, commercial, manufacturing, and transportation. The RECS is used to collect data from the residential sector. The Energy Policy Act (EPACT) of 1992 directs EIA to conduct a residential survey at least once every 3 years.

The RECS provides information on the use of energy in residential housing units in the United States. This information includes the physical characteristics of the units, the appliances utilized, the occupants, the types of fuels being used, the amount of energy used, and other energy use characteristics.

Because the RECS cannot survey every household in the Nation, a statistical sample representing all U.S. households is chosen. The RECS sample design was initially developed during 1979 and 1980, and was used in the 1980, 1981, and 1982 RECS. The sample design was updated in 1984 and 1993 using the results of the 1980 and 1990 Censuses, respectively. This report, therefore, distinguishes between the three basic sample designs (1980, 1984, and 1993), and the nine RECS fieldings (1978, 1979, 1980, 1981, 1982, 1984, 1987, 1990, and 1993).

The RECS sample design has always been a national multistage area-probability cluster-sample design. The basic sampling principles followed are in accord with recommendations in standard sampling texts, which evolved from sampling theory.[1] Many of the national surveys of households or housing units that are conducted by the Federal government use the same type of design as used for RECS. In particular, the design for RECS and the design for the Current Population Survey (CPS) are based on the same sampling principles.[2]

---

[1]Cochran, W.G.: *Sampling Techniques*, 3rd Ed. New York, NY, John Wiley & Sons, 1977; Hansen, M.H., Hurwitz, W.N. and Madow, W.G.: *Sample Survey Methods and Theory*, Vols. I and II, New York, NY, John Wiley & Sons, 1953; Kish. L.: *Survey Sampling*, New York, NY, John Wiley & Sons, 1965.

[2] The Technical Paper 40, "The Current Population Survey Design and Methodology," Department of Commerce, Bureau of the Census, January 1978.

With this type of multistage area-probability cluster-sample design every occupied housing unit in the country has a known chance of being selected for RECS, yet the design does not require the use of a comprehensive up-to-date list of all housing units in the country. Such a list would be very costly to develop and maintain. (According to the 1990 Census, there were 91,947,410 occupied housing units in the country.) The Bureau of the Census and the U.S. Postal Service are currently developing such a list as part of the planning for the 2000 Decennial Census. The Bureau of the Census plans to maintain and update the list. Consequently, the list will be available for use in the development of the sample design for post-2000 household surveys conducted by the Bureau of the Census. EIA is hopeful that the next major revision of the RECS sample design (following the 2000 Census) can utilize this valuable resource.

With a multistage area-probability cluster-sample design, the interviews are geographically clustered. The clustering reduces both design and interviewing costs, though designs with less clustering would result in more precise estimates. For each stage of the design, the procedures for (1) defining the sampling units, (2) stratifying the sampling units, and (3) selecting the sampling units, are chosen in an attempt to balance between reducing survey costs and meeting the precision requirements for national, regional, and Census division data.

The dual objectives of reducing survey costs and minimizing the variance of survey estimates (maximizing the accuracy of the estimates), pull in opposite directions. The more the interviews are clustered, the lower the survey costs. On the other hand, for a fixed sample size, the more the interviews are clustered, the larger the variance of the survey estimates. Procedures for optimizing a survey design to balance the two objectives were used in the development of the RECS sample design[3].

History

The RECS has had four different sample designs since the survey began in 1978. This report will focus on the four designs, and the individual surveys will be grouped by their sample design year, as presented in the following sections.

National Interim Energy Consumption Survey and the Household Screener Surveys

Interim versions of RECS were conducted in 1978 and 1979. The 1978 survey was called the National Interim Energy Consumption Survey (NIECS), and the 1979 survey was called Household Screener Survey (Screener). The sample design for the NIECS and the Screener was the same, a multistage area-probability cluster-sample, which was the property of the survey contractor. The NIECS used a multi-purpose sample design, one not specifically developed to collect energy-related data, and which had been used in non-energy-related surveys. It covered the 48 contiguous States and the District of Columbia, but did not cover Hawaii and Alaska. The NIECS supported estimates at the Census region level, but not at the Census division level.

The 1980 Sample Design (1980, 1981, and 1982 RECS)

The sample design for the RECS was expressly developed to collect energy-related data in the residential sector. The RECS design is the property of EIA and is used only for RECS or other EIA energy-related surveys. This design covered all 50 States and the District of Columbia, and supported estimates at the Census division level. Other EIA surveys based on the RECS design include Residential Transportation Energy Consumption Survey (RTECS) and Commercial Buildings Energy Consumption Survey (CBECS). A brief description of these designs will follow later in this section on the history of RECS.

[3] Hansen, Hurwitz, and Madow, *Sample Survey Methods and Theory Volume I Methods and Applications*, John Wiley & Sons, 1953.

The 1984 Sample Design (1984, 1987, and 1990 RECS)

In addition to the 1980 Census results, the 1984 design incorporated changes in the definitions of Metropolitan Statistical Areas (MSA's), and changes in the priorities of the survey. The sample design for the 1984 RECS also contained a longitudinal panel and an incoming panel. The longitudinal panel was taken from the 1980 sample design, and the incoming panel of sample households was taken from the 1984 design. (See Chapter 7 for more details on longitudinal and incoming panels.) The 1987 RECS was the first to employ the 1984 sample design for all observations. This design was also used in the 1990 RECS.

The 1993 Sample Design (1993, 1996, and 1999 RECS)

The 1993 revision of the RECS sample design incorporated changes in the definitions of MSA's, and improvements in the stratification of the sampling units in all stages. In developing this sample design, additional energy-related characteristics of the population (such as estimated energy expenditures) were used in the stratification and selection procedures. The primary stage for the 1993 RECS used the new design. However, while the initial plan was to completely redevelop all three stages of the sample design for the 1993 RECS, in order to lower costs for some primary-stage units, their respective secondary stages were carried over from the 1984 design. (See Chapter 6 for more details.)

Current plans call for the 1996 RECS to use the full 1993 sample design. All secondary-stage units to be used for the 1996 RECS will have been selected during the 1993 design effort. No secondary-stage units will be carried over from the 1984 design to the 1996 RECS. The 1993 design will also be the basis for the 1999 RECS.

Residential Transportation Energy Consumption Survey (RTECS)

The sample design for the RTECS involves selecting a subsample of the respondents from the RECS of the previous year. Therefore, any change in the sample design for RECS automatically applies to the RTECS.

Commercial Buildings Energy Consumption Survey (CBECS)

The primary stage for the 1984 RECS sample design was also used as the primary stage in the 1986 CBECS sample design. However, the secondary stage and the elementary-unit stage in the 1986 CBECS sample design were totally different than those used in the 1984 RECS. The CBECS sample design will be updated prior to its use in 1995. During this redesign effort, the decision on the feasibility of using the same first-stage design for both RECS and CBECS will be reviewed.

Organization of this Report

The RECS design is not static, it changes as priorities change. The following chapters present summaries of the details and changes in the sample design throughout the development of the RECS.

Chapter 1 gives a brief overview of the RECS. Chapter 2 gives a summary of the multistage area-probability design that is used for the RECS. Chapter 3 discusses the objectives of the RECS sample design, and the constraints under which it was developed. Chapter 4 discusses the optimization analyses that were performed prior to the 1980, 1984, and 1993 RECS. Chapters 5 and 6 present details on the primary and secondary

stages of the sample, respectively. Chapter 7 presents details on the elementary-unit stage, and discusses additional features of the RECS design, including the longitudinal component, the low-income supplement, the new construction supplement, and the new construction update procedures. Chapter 8 presents a discussion of the reasons for periodically revising the RECS sample design. Appendix A shows the nine Census divisions and the ten Federal regions and maps. Appendix B presents detailed instructions to field workers on the housing unit definition. Appendix C presents the strata listings by Census division.

2. Multistage Area-Probability Design

It is necessary to select a sample of households for the RECS because the cost and burden of interviewing every household in the United States would be prohibitive. The RECS sample design has always been a national multistage area-probability cluster-sample design. This type of design is used for many of the national surveys of households or housing units that are conducted by the Federal government. The following is a description of the multistage area-probability design with specific details about the RECS design.

In each RECS cycle, the sample design is divided into a primary stage, a secondary stage, and an elementary-unit stage. These three stages are briefly outlined in the box below.

Stages of RECS Sample Design

- **Primary Stage:** In this stage, ***Primary Sampling Units*** (PSU's) are defined, grouped into strata, and selected for the design. The procedures in this stage can be divided into three groups:

  1. ***PSU Formulation Procedures***: Counties, independent cities, and other county-like geographic areas are used to form PSU's.

  2. ***Strata Formulation Procedures***: The PSU's are grouped into strata.

  3. ***PSU Selection Procedures***: One PSU is selected from each stratum.

- **Secondary Stage:** In this stage, the PSU's that were selected in the primary stage are divided into ***Secondary Sampling Units*** (SSU's). Multiple SSU's are selected from each PSU, and those selected are divided into ***Listing Segments***, with one listing segment selected per SSU. The procedures in this stage can be divided into five groups:

  1. ***SSU Formulation Procedures***: Census geographic units such as Census tracts, enumeration districts, block groups, and individual blocks are used to form SSU's.

  2. ***SSU Selection Procedures***: The SSU's are ordered or stratified. The selection procedures guarantee that these sample SSU's reflect the distribution of the values of the variates used in the ordering or stratification.

  3. ***Rough-Counting Procedures***: For the larger SSU's selected in the secondary stage, field workers canvass the SSU by automobile or on foot, recording gross estimates of the number of housing units on each block face.

  4. ***Segment Formulation Procedures***: The SSU's selected in the secondary stage are divided into ***listing segments*** using the rough counts of housing units. Smaller SSU's are defined as single listing segments.

  5. ***Segment Selection Procedures***: A single listing segment is selected for each SSU chosen in the secondary stage.

Stages of RECS Sample Design (Continued)

- **Elementary-Unit Stage:** Housing units make up the most elementary units in the RECS. In this stage, housing units are selected from their respective listing segments and given to the field workers for their interviewing assignments. The elementary-unit stage can be subdivided into three groups of procedures:

  1. ***Field-Listing Procedures***: For each listing segment selected, field workers canvassed the segment on foot, identifying and listing the address of all housing units. The lists prepared by the field worker are used by central office workers to develop a comprehensive list of housing units in the segment.

  2. ***Penultimate Cluster Selection Procedures***: From each listing segment a single ***penultimate cluster*** is selected. If the listing segment is not too large, the penultimate cluster will equal the listing segment. If the listing segment contains too many housing units, the penultimate cluster will be made up of a subset of these units.

  3. ***Interviewer Assignment Procedures***: An ***ultimate cluster*** of approximately five housing units is selected from each penultimate cluster. The housing units in the ultimate clusters make up the sample in which the field workers conduct their occupant interviews.

It may be argued that the penultimate clusters are the actual secondary-stage units. If this is true, then it follows that the field-listing procedures and the penultimate cluster selection procedures are really part of the secondary stage. However, EIA has chosen to place the field-listing procedures and the penultimate cluster selection procedures in the elementary-unit stage, because these procedures deal with the elementary units.

Design Considerations

The number of PSU's selected in the primary stage, the number of SSU's (or listing segments) per PSU, and the number of observations in the ultimate cluster, are chosen according to survey design principles. These decisions are made with the two-fold goal of reducing survey costs, and minimizing the variance of survey estimates.

The first two stages in the design (primary and secondary), are used to select a sample of listing segments.

The size of the listing segments is determined by two goals:

1. The approximate number of housing units in a listing segment should be small enough so that a comprehensive list of these units can be developed by field workers at a reasonable cost

2. A listing segment needs to contain a sufficient number of housing units to cover the needs of several cycles of the RECS survey.

During the initial RECS design effort, as well as during the 1984 design update, the minimum size of a listing segment was 25 housing units. For the 1993 design, the minimum size was expanded to 50 housing units for the core sample, and 96 housing units for SSU's in a new construction supplement. The minimum size for the core sample SSU's was increased to give more flexibility in the elementary-unit stage not only for the 1993

RECS, but also for the 1996 and 1999 RECS. The 1980 and 1984 designs called for the listing segments and penultimate clusters to be periodically replaced. If a 1993 RECS SSU does not experience substantial new residential construction, current plans call for using the corresponding 1993 penultimate cluster in the 1996, the 1999, and perhaps in the 2002 RECS. The expanded minimum size for the penultimate clusters ensures that they will contain enough housing units for four RECS cycles, as well as provide for an oversampling of low-income areas and new housing units. The expanded minimum size also allows for an increase in the core sample size for future RECS. The minimum size for SSU's in the new construction supplement is larger than that of core sample SSU's, due to the higher sampling rate for new housing units.

The penultimate cluster selection procedures are incorporated into the design as a cost-saving feature. Only the housing units in the penultimate cluster are keyed into the data base used to select the ultimate cluster. In addition, apart from major design revisions, or the occurrence of significant new construction (see Chapter 7), the list of housing units is updated only for the compact area covered by the penultimate cluster.

The most expensive parts of the design effort are the rough-counting procedures of the secondary-unit stage, and the field-listing procedures of the elementary-unit stage. Both of these procedures require the use of field workers in order to obtain data, and central office staff to edit, clean, and process the data. In the rough-counting procedures, field workers are used to obtain an approximate number of housing units by block face in the selected SSU's. In the field-listing procedures, field workers are used to prepare the comprehensive lists of housing units in the selected listing segments. The entire primary stage, and the SSU formulation and selection procedures of the secondary stage, can be done in a central office using maps, computer software, and data bases obtained from vendors and from the Bureau of the Census. Similarly, the segment formulation and selection procedures, the penultimate cluster selection procedures, and the interviewer assignment procedures, can be done in a central office using the rough counts and comprehensive lists prepared by the field workers.

If a SSU is small enough to economically list all of its housing units, then the rough-counting procedures, the segment formulation procedures, and the segment selection procedures are not needed. In this case, the SSU is not divided into listing segments. The entire SSU is defined as a single listing segment and is automatically selected.

The target number of housing units in a SSU for the 1980 and 1984 designs was 400. All of the SSU's for these designs, therefore, were much larger than the minimum size. Consequently, the rough-counting procedures, the segment formulation procedures, and the segment selection procedures were necessary for all of the 1980 and 1984 SSU's.

In the 1993 design, the target number of housing units was reduced to a quantity that was much closer to the minimum size. In the 1993 RECS, the core SSU's were defined so that their projected minimum number of housing units was 50, and the new construction supplement SSU's were defined so that their projected minimum number of housing units was 96. As a result, the proportion of SSU's in which the rough-counting procedures, the segment formulation procedures, and the segment selection procedures were necessary, was much smaller. This resulted in a cost savings for the design phase.

In summary, the number of cases where rough counting was necessary was substantially reduced in the 1993 design. This was possible because the minimum size for a listing segment was increased for the 1993 design effort, while the size of the SSU's was reduced. Prior to the 1990 Census, the Bureau of the Census divided the entire country into Census tracts and divided all Census tracts into blocks. This enabled the 1993 survey design team to use smaller Census units to define SSU's in rural areas. In the 1980 Census, the smallest Census unit in many rural areas was an enumeration district.

3. Objectives and Constraints to the RECS Sample Design

The RECS was designed to collect data on energy-related characteristics, both of housing units and the households living in these units. Consequently, the RECS elementary-sampling units were occupied housing units. The RECS was further restricted to housing units that are the primary residence of the occupants. As a result, there is a one-to-one correspondence between households and occupied primary-housing units. Since the elementary-sampling units for the CPS are households, its design was used as a model for the development of the sample design for RECS.

The CPS is mainly concerned with demographic and labor-force statistics, while the RECS is concerned with energy characteristics. Also, while the sample size for the CPS is approximately 60,000 observations, the sample size for RECS is approximately 5,000 observations. Because of these variations, there are differences between the sample designs for RECS, and the design for the CPS. For example, the number of PSU's and SSU's used for RECS will be smaller than the number used for the CPS. In addition, the variables used to stratify or order the sampling units for the three stages of the sample design will not be the same for RECS as for the CPS.

1980 Sample Design

The 1980 RECS sample design was limited by the following constraints[4]:

- Data are to be statistically reliable for nine Census divisions and ten Federal regions. (These are discussed in Chapter 5.)
- A self-weighting national sample should be included as the core of the total design; the balance of the sample is to be allocated so that, as a minimum, a specified level of precision is achieved for each Census division and Federal region
- Sample households are to be selected from as many States as possible; inclusion of some sample units in Alaska and Hawaii is to be assured
- In addition to regional geographic factors, stratification modes should be based, to the extent possible, on principal home heating fuel and climatological factors
- As a quality control factor, a minimum of two interviewers should be available for data-collection activity within reasonable travel distance of each cluster of sample households.

These additional constraints also affected the design:

- The design must provide for the contingency that RECS may have to be fielded every 6 months
- The design must provide for the contingency of including a longitudinal component in future RECS
- The 1980 Census population data, and data on the number of households, were not available during the design phase. The 1970 Census data were available, but were outdated for some localities

[4] Residential Energy Consumption Survey, *Sample Design and Methodology National Household Surveys: 1980, 1981, and 1982*, Response Analysis Corporation, Princeton, New Jersey, August 1983.

- The most recent definitions of Metropolitan Statistical Areas--issued by the Office of Management and Budget formerly known as Standard Metropolitan Statistical Areas--were those as of 1980, and developed prior to the completion of the 1980 Census
- The 1980 RECS was designed to produce 5,000 completed observations
- The design must provide for the contingency that would extend the survey to produce State-level data.

1984 Sample Design

In the 1984 RECS redesign effort, the following changes were made in the design constraints:

- The mandate requiring the 1984 RECS to provide statistically reliable estimates for each of the ten Federal regions was eliminated. As a result, the number of strata used in the first stage of the design was reduced. The mandate requiring the 1984 RECS to provide statistically reliable estimates for each of the 9 Census divisions was observed
- The 1980 Census data would be available for the design effort
- The June 1983 definitions of MSA's (developed following the 1980 Census) would be available for the design effort.

1993 Sample Design

In the 1993 RECS redesign effort, the following changes were made in the design constraints:

- The 1990 Census data would be available for the design effort
- The June 1990 definitions of MSA's would be available for the design effort (developed prior to the completion of the 1990 Census)
- The design was required to have the capability to oversample newly constructed housing units.

The following changes were made to the RECS sample design as a result of the 1993 redesign effort:

- The number of strata of PSU's was reduced primarily because of larger increases in the PSU-level survey costs (for example, administrative costs, and the cost of recruiting and training interviewers), in contrast to SSU-level survey costs and the elementary-unit level costs, which showed smaller increases
- The stratification procedures used in the primary stage were changed in order to use more energy-related characteristics
- The SSU's for metropolitan PSU's were stratified by energy-related characteristics and geography. For nonmetropolitan PSU's, the SSU's were stratified by geography alone
- A new construction supplement was incorporated into the sample design. Part of the supplement was the selection of 150 SSU's, which were expected to include high percentages of new homes
- The number of PSU's and SSU's assigned to each Census division was altered to more closely reflect the relative 1990 Census population of each Census division.

4. Optimization Analysis

Prior to each major revision in the RECS sample design, an optimization analysis is conducted to determine the optimum number of PSU's, SSU's per PSU, and observations per SSU. The results of this analysis are used to guide decisions concerning the number of strata in the primary stage, the number of SSU's per PSU in the secondary stage, and the size of the ultimate clusters in the elementary-unit stage.

In each case, the optimization analysis only considered the effect of the sample design on the precision of national level statistics. It would be possible to expand the optimization analysis to cover the effect for subnational level statistics, but the vast number of subnational level statistics that are of interest implies that it would be very difficult and time consuming to conduct an optimization analysis that covers the effect on all statistics of interest. Consequently, each optimization analysis was restricted to covering the effect on the precision of national level statistics. Hence; the results of the optimization analysis are not completely followed in determining the sample design parameters. The results are used as a guide for determining the design parameters.

In a multi-stage sample design, the level of precision is affected not only by the total number of sample households, but also by the interaction of all the design features: the number of PSU's, the number of SSU's, the average size of an ultimate cluster, the "within" versus the "between" PSU variance, and the effectiveness of stratification, etc. The interaction between the level of precision and the design factors is approximated by the following formula[5]:

$$V_r^2 \doteq V^2 [1/m\bar{n}\bar{q}][\delta_1\bar{n}\bar{q} + 1 + \delta_2(\bar{q} - 1)] \quad (1)$$

where:

- V_r^2 is the relvariance of an estimated mean or percentage derived from the survey results
- V^2 is the unit relvariance in the universe
- m is the number of sample PSU's
- $\bar{n}$ is the expected value of the average number of SSU's per sample PSU
- $\bar{q}$ is the expected value of the average number of elementary units in an ultimate cluster that result in completed interviews
- δ_1 is the "within" PSU measure of homogeneity
- δ_2 is the "within" SSU measure of homogeneity.

Note that the product of m, $\bar{n}$, and $\bar{q}$ is the total sample size.

The values of V^2, δ_1, and δ_2 are estimated using survey data. For the optimization conducted prior to the development of the 1993 design, 1990 RECS data were used to estimate the values for these parameters.

[5] Hansen, Hurwitz, and Madow, *Sample Survey Methods and Theory Volume I Methods and Applications*, p 403, John Wiley & Sons, 1953.

In order to determine the optimum set of values for m, $\overline{n}$, and $\overline{q}$, where the total sample size is held fixed, it is necessary to approximate the survey costs (i.e., design, data collection, and data processing costs) as a function of m, $\overline{n}$, and $\overline{q}$. The following simple cost function was used to approximate the relationship[6]:

$$C = C_1 m + C_2 m\overline{n} + C_4 m\overline{n}\,\overline{q} \qquad (2)$$

where: C is the portion of the total survey cost associated with "volume related" activities

C_1 is the cost per selected PSU

C_2 is the cost per SSU

C_4 is the cost per completed interview.

If the sole purpose of RECS was to produce a national estimate of a single statistic, such as the number of housing units in the country using electricity as a primary heating source, or the average electricity consumption per housing unit, then the design should agree closely with the results of the optimization analysis. However, since RECS is a multi-purpose survey, the data obtained through it will be used to produce estimates for many housing-unit characteristics. In addition, EIA is interested in producing not only national estimates, but also estimates by Census division, by housing-unit type, by household income, as well as many other categories of residential energy use. As a consequence of the multipurpose nature of RECS, the value of its design varies depending on the characteristic of interest. This particular design will produce more precise estimates for some characteristics and less for others.

The optimization analysis in the 1993 design is conducted on 1990 RECS data, as this represents the latest data available. Figure 1 defines the parameters used in the optimization analysis, and Table 1 lists these parameters along with their resulting optimal values. The cost per primary stage unit (C_1) is $3,095.50, the secondary stage cost per unit (C_2) is $368.50, and the cost per observational unit (C_4) is $184.00. The optimization analysis was based on the assumption that the core sample size would be 5,095 observations. If the optimization analysis had been done with a fixed cost (that varied from the final effective cost) instead of a fixed sample-size, it would have called for a different value of m.

In an attempt to generalize the specific results contained in Table 1, the twenty-three ratio estimates were grouped into four classes: (1) consumption and expenditure variables; (2) housing unit characteristic variables; (3) variables relating to appliance ownership; and (4) demographic characteristics of home owners. Within each of these four classes, average values of intraclass correlations were computed. The class averages for the intra-class correlation statistics were then used to compute optimal allocation solutions for each general class of variables. Table 2 provides the optimal allocation results for the resulting grouped computations.

The values that were used for the 1993 RECS sample design were m = 116, $\overline{n}$ = 12.6, $\overline{q}$ = 4.11, and the core sample size was equal to 6,000 observations. Chapters 5 and 6 will discuss in detail, the number of PSU's and the number of SSU's that were actually used in the 1993 RECS sample design.

The number of PSU's (m) used in the 1993 RECS sample design does fall within the range indicated by the optimization analysis. However, the average number of SSU's per PSU ($\overline{n}$) used, and the average number of observations per ultimate cluster ($\overline{q}$), differ substantially from that indicated by the optimization analysis. In particular, the average number of SSU's per PSU is smaller than that indicated by the optimization analysis, while the average number of observations per ultimate cluster is larger.

[6] Hansen, Hurwitz, and Madow, *Sample Survey Methods and Theory Volume I Methods and Applications*, p 408, John Wiley & Sons, 1953.

Figure 1. Description of Ratio Statistics Used in the 1990 RECS Components of Variance Analysis

Ratio Statistic Name	Description
Btu Electricity	Mean Electricity Consumption per Household in Btu
Btu Natural Gas	Mean Natural Gas Consumption per Household in Btu
Dollars Electricity	Mean Annual Household Expenditure For Electricity in Dollars
Dollars Natural Gas	Mean Annual Household Expenditure For Natural Gas in Dollars
Total Btu	Mean Total Energy Consumption per Household in Btu
Total Dollars	Mean Total Energy Expenditure per Household in Dollars
Awnings	Proportion of Housing Units with Window Awnings
Bedrooms	Mean Number of Bedrooms per Household
Blinds/Drapes	Proportion of Housing Units with Window Blinds or Drapes
Clothes Washer	Proportion of Households with a Clothes Washer
Dishwasher	Proportion of Households with a Dishwasher
Drivers in HH	Mean Number of Licensed Drivers per Household
Electric Dryer	Proportion of Households with an Electric Clothes Dryer
HH Sex	Proportion of Households with a Male Head
HomeArea	Mean Number of Square Feet of Space per Household
Income $35,000+	Proportion of Households with Annual Income Over $35,000
HU Tenure	Proportion of Households Owned by the Current Resident
No. Complete Baths	Mean Number of Bathrooms per Housing Unit
Refrigerator Size	Mean Cubic Feet of Refrigerator Storage Space
Stories	Mean Number of Stories per Housing Unit
Color TV	Mean Number of Color Televisions per Household
Vehicles	Mean Number of Cars and Trucks per Household
Windows	Mean Square Feet of Window Space per Household

Source: Energy Information Administration, *Analysis of Components of Variance for the Residential Energy Consumption Survey* (Draft), Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), December 16, 1993.

If the 1993 RECS sample design closely followed the optimization analysis, the average number of observations per ultimate cluster for the core sample would have equaled approximately 1.5. In order to have 6,000 observations, the total number of SSU's ($m \times \overline{n}$) would equal approximately 4,000. To use such a large number of SSU's would have substantially increased the design and survey costs. In order to stay within budget, fewer than 4,000 SSU's were used.

If $\overline{q}$ and $\overline{n}$ are within the optimal range, and m is smaller than 116, then the total number of sample SSU's would be reduced, while the total core sample size would also be reduced. A design using such values for $\overline{q}$,

$\bar{n}$, and m would result in more precise national estimates of energy statistics than the design that was actually used, but the precision for the corresponding estimates for many subsections of the population would have been lowered.

Table 1. Optimum Multistage Allocation-Solution Estimates for 1990 RECS Ratio Statistics

Variable	δ_1	δ_2	Optimal Allocation		
			$\bar{q}$	$\bar{n}$	m
Btu Electricity	.01241	.37094	1.84	15.84	174.5
Btu Natural Gas	.02278	.56636	1.24	14.45	284.7
Dollars Electricity	.01308	.33392	2.00	14.64	174.1
Dollars Natural Gas	.01968	.56296	1.25	15.50	263.6
Total Btu	.01271	.56358	1.25	19.30	212.0
Total Dollars	.00287	.41881	1.67	35.02	87.3
Awnings	.00036	.10751	4.08	50.10	24.9
Bedrooms	.01141	.69468	0.94	22.62	240.1
Blinds/Drapes	.00075	.07600	4.93	29.08	35.5
Clothes Washer	.00342	.73611	0.85	42.52	141.4
Dishwasher	.00153	.25621	2.41	37.45	56.4
Drivers in HH	.00320	.27615	2.29	26.93	82.6
Electric Dryer	.00660	.44753	1.57	23.86	135.8
HH Sex	.00132	.06892	5.20	20.97	46.7
Homearea	.01418	.68448	0.96	20.13	263.4
Income $35,000+	.00011	.47697	1.48	18.64	138.9
HU Tenure	.00899	.65739	1.02	24.78	201.3
No. Complete Baths	.00650	.55433	1.27	26.78	150.0
Refrigerator Size	.00394	.34716	1.94	27.22	96.5
Stories	.01403	.69230	0.94	20.36	265.2
Color TV	.00098	.26064	2.38	47.38	45.1
Vehicles	.00445	.36352	1.87	26.20	103.8
Windows	.00675	.62120	1.11	27.80	165.9

Source: *Analysis of Components of Variance for the Residential Energy Consumption Survey*, Sampling Section, Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), December 16, 1993.

Table 2. Optimum Multistage Sample Allocation Based on Average Values of Estimates of δ_1, and δ_2

Variable	δ_1	δ_2	Optimal Allocation		
			$\bar{q}$	$\bar{n}$	m
Consumption and Expenditure Variables	.013922	.469427	1.50	16.83	201.2
Housing-Unit Characteristic Variables	.007711	.490072	1.44	23.10	152.8
Appliance Variables	.003444	.383903	1.79	30.60	92.9
Demographic Variables	.003474	.401091	1.73	31.14	94.6

Source: *Analysis of Components of Variance for the Residential Energy Consumption Survey*, Sampling Section, Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), December 16, 1993.

Even though the values of $\overline{q}$, $\overline{n}$, and m that were used in the 1993 RECS design differ from those called for by the optimization analysis, the precision of the national estimates remains near the optimum (due to the breadth of the optimum), and the precision of most subnational estimates of interest remains greater than would those yielded by a design following the results of the optimization analysis.

If the population of households is divided into three income categories (such as low, middle, and high), then the effective size of $\overline{q}$ for these categories would be approximately one-third of their original value; hence the value of $\overline{q}$ used for the 1993 design would fall in the range called for by the optimization analysis.

The 1980 RECS sample design used 131 PSU's, and an average of 11.6 SSU's per PSU. The 1984 RECS sample design used 129 PSU's, and an average of 11.8 SSU's per PSU. The number of observations per SSU averaged 3.3 for the 1980 through 1990 RECS.

The 1993 RECS sample design resulted in similar values. The number of PSU's was reduced from 129 in 1984 to 116 in 1993, while the number of SSU's in the core sample was reduced from 1,516 in 1984 to 1,460 in 1993. The number of observations per SSU was increased from an average of 3.3 to 4.1 for the core sample of the 1993 RECS. The reduction in the number of PSU's was due to cost increases at the PSU-level which outpaced those at the SSU and the elementary-unit levels. The number of observations per SSU was increased in order to provide a larger sample size, at the most cost-effective level.

There would have been a greater reduction in the number of PSU's if the 1993 RECS design had not required the survey to support Census division-level estimates with a specified maximum Relative Standard Error (RSE), as is discussed below. For this reason, the minimum number of PSU's in a Census division was set at 8.

One of the constraints in the 1993 RECS sample design was that the RSE of the estimate for the average energy expenditures per household would be less than or equal to the predetermined levels. The RSE for the National level estimates would be a maximum of 1.25 percent. The estimates for each Census region would have an RSE of no more than 2.75 percent, while the estimates for each Census division would have an RSE of 4.5 percent, or less. These constraints were determined by examining the results of the 1980 through 1990 RECS. Simulations projected that the sample design with the above values for $\overline{q}$, $\overline{n}$, and m would result in RSE estimates below the constrained maximum RSE values listed above.

5. Primary Stage of the RECS Sample Design[7]

During the primary stage, the PSU's are defined, grouped into strata, and selected for the design. The following decisions are made during the design and implementation of the primary stage:

- Whether to select just one PSU per stratum, or to select two or more PSU's per stratum
- How will Census division boundaries and Federal region boundaries constrain the sample design
- The total number of strata to use in the primary stage
- Which criteria to use when forming PSU's
- Which criteria to use when grouping PSU's into strata
- How to select PSU's from strata?

Each of these decisions can be divided into one of the following three procedures: (1) PSU formulation procedures, (2) strata formulation procedures, and (3) PSU selection procedures. These decisions will be discussed in detail in the following sections.

Number of Primary Sampling Units' per Stratum

There are two basic competing strategies for the primary stage of a multistage area-probability sample design:

1. One PSU per stratum strategy: Stratify the sampling units as much as possible, then select only one sampling unit per stratum
2. Two or more PSU's per stratum strategy: Restrict the amount of stratification and if a stratum contains more than one sampling unit, select at least two of them.

For either strategy, an optimization analysis is used to determine the approximate number of sampling units that will be selected during each stage of the design.

The main advantage of the one-PSU per stratum strategy is the use of stratification to the fullest extent in order to improve the accuracy of the survey estimates. Its main disadvantage to survey estimates is the inability to obtain a pure estimate of the sampling errors. Stratifying to the fullest extent may improve the estimate, but as a result, it is necessary to use a variance estimation technique that does not utilize a pure estimate of the between sampling unit variance.

The main advantage of the two-PSU's per stratum strategy is that it is possible to obtain a pure estimate of the variance using a two-sampling unit per stratum design. Its main disadvantage is that fewer strata are used, and as a result, the precision of the estimate is lowered.

An early decision was made to use the one-PSU per stratum strategy for the primary stage. The sample design for all cycles of RECS have used the one-PSU per stratum strategy, and the 1993 sample design will continue to use this strategy. The only exception, in the 1993 sample design, was that the Fort Worth, Texas MSA; the

[7]*Residential Energy Consumption Survey Primary Stage Sample Design Plan* (Third Draft), Sampling Section, Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), February 1, 1993.

San Antonio, Texas MSA; and the Austin, Texas MSA were placed in the same stratum and two MSA's were selected from it. The reasons for using a double strata in this case are discussed in detail in the section titled, "Number of Strata by Type and Census Division" in this chapter.

Census Divisions and Federal Regions

As mentioned earlier, the 1980 RECS was designed in such a way that its data could be used to estimate energy characteristics for each of the nine Census divisions and each of the ten Federal regions. This prevented PSU's, and the strata used in the primary stage, from crossing the Census division or Federal region boundaries. However, because the 1984 and 1993 RECS designs were no longer required to produce estimates for the ten Federal regions, the PSU and strata boundaries could now cross Federal region boundaries (but they were still unable to cross Census division boundaries). Figure 2 lists the nine Census divisions and ten Federal regions. They are also illustrated in map form in Appendix A.

Figure 2. List of the States Within Census Divisions and Federal Regions

The nine Census divisions are defined as follows:

1. New England:	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont
2. Middle Atlantic:	New Jersey, New York, and Pennsylvania
3. East North Central:	Illinois, Indiana, Michigan, Ohio, and Wisconsin
4. West North Central:	Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota
5. South Atlantic:	Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, and West Virginia
6. East South Central:	Alabama, Kentucky, Mississippi, and Tennessee
7. West South Central:	Arkansas, Louisiana, Oklahoma, and Texas
8. Mountain:	Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming
9. Pacific:	Alaska, California, Hawaii, Oregon, and Washington.

The 10 Federal Regions are defined as follows:

Federal Region 1:	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont
Federal Region 2:	New Jersey and New York
Federal Region 3:	Delaware, District of Columbia, Maryland, Pennsylvania, Virginia, and West Virginia
Federal Region 4:	Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, and Tennessee
Federal Region 5:	Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin
Federal Region 6:	Arkansas, Louisiana, New Mexico, Oklahoma, and Texas
Federal Region 7:	Iowa, Kansas, Missouri, and Nebraska
Federal Region 8:	Colorado, Montana, North Dakota, South Dakota, Utah, and Wyoming
Federal Region 9:	Arizona, California, Hawaii, and Nevada
Federal Region 10:	Alaska, Idaho, Oregon, and Washington

The New England Census division and the 1st Federal region are identical. For all other Census divisions, there does not exist an identical Federal region. Similarly, for all other Federal regions, there does not exist an identical Census division. The combined boundaries of the nine Census divisions and the ten Federal regions produce 17 intersections. These intersections are listed in Figure 3. The strata boundaries for the 1980 RECS design could not cross the boundaries of the 17 intersections.

Figure 3. State Intersections of Census Divisions and Federal Regions

Federal Region	Census Division	Intersection of Federal Region and Census Division	
1	New England	1-	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont
2	Middle Atlantic	2-	New Jersey and New York
3	Middle Atlantic	3-	Pennsylvania
	South Atlantic	4-	Delaware, District of Columbia, Maryland, Virginia, and West Virginia
4	South Atlantic	5-	Florida, Georgia, North Carolina, and South Carolina
	East South Central	6-	Alabama, Kentucky, Mississippi, and Tennessee
5	East North Central	7-	Illinois, Indiana, Michigan, Ohio, and Wisconsin
	West North Central	8-	Minnesota
6	West South Central	9-	Arkansas, Louisiana, Oklahoma, and Texas
	Mountain	10-	Mew Mexico
7	West North Central	11-	Iowa, Kansas, Missouri, and Nebraska
8	West North Central	12-	North Dakota and South Dakota
	Mountain	13-	Colorado, Montana, Utah, and Wyoming
9	Mountain	14-	Arizona and Nevada
	Pacific	15-	California and Hawaii
10	Mountain	16-	Idaho
	Pacific	17-	Alaska, Oregon, and Washington

Source: Energy Information Administration, Residential Energy Consumption Survey, *Sample Design and Methodology National Household Surveys: 1980 - 1981, and 1982*, Response Analysis Corporation (Princeton, NJ), August 1983.

Number of Strata to Use in the Primary Stage

The approximate total number of primary stage strata that will be used is determined during the optimization analysis. The final number of strata reflects decisions made concerning the number of strata per Census division, and within Census divisions, per strata type (e.g., large metropolitan areas, small metropolitan areas, non-metropolitan areas, and mixed areas). For each Census division, including the national total, Table 3 is divided into two parts (a and b), which provides the following data:

- The number of primary stage strata for the 1980, 1984, and 1993 sample designs
- The population: the 1980 design employs the 1978 population estimates, the 1984 design employs the 1980 Census results, and the 1993 design employs the 1990 Census results
- The average population per primary stage stratum: the average population given in the 1980 design is the average 1978 population, in the 1984 design it is the average 1980 population, and in the 1993 design it is the average 1990 population
- The 1990 Census number of households
- The average number of households per stratum for the 1993 sample design.

The reason for determining the number of households and the average number of households per stratum, is because the Measure of Size (MOS) used to select the PSU's changed from PSU population in the 1980 and 1984 designs, to PSU number of households in the 1993 design. (The MOS is discussed later in this chapter.) From a practical perspective, each of the two variables used for the MOS is equally acceptable in devising a multipurpose probability-sampling design.

Because the RECS sample design has always used a one-PSU per stratum approach, there is a one-to-one correspondence between the strata and the PSU's selected. Therefore, when discussing the design, there is a tendency to interchange the terms stratum and PSU. To illustrate this, the 1980 sample design used 131 PSU's; this means that the sample design for the 1980 RECS grouped the PSU's into 131 strata and selected one PSU per stratum, which yielded 131 PSU's.

The 1980, 1984, and 1993 sample designs each determined the set of self-representing PSU's by a standard sampling approach; the PSU's that exceeded in MOS a substantial fraction of the average MOS per stratum were identified as the self-representing set. Therefore, in each of the three sample designs, the New York counties of the New York Primary Metropolitan Statistical Area were grouped into two PSU's. New York City (Bronx, Kings, New York, Queens, and Richmond counties), and its northern suburbs (Putnam, Rockland, and Westchester Counties), formed one PSU, which was placed into a stratum by itself. This PSU contained 3,252,000 households in 1990, larger than any other PSU. The Long Island counties (Nassau and Suffolk) formed another PSU that was placed in a stratum by itself. The Illinois region of the Chicago PMSA, as well as Los Angeles county, were also single PSU's. These three largest self-representing PSU's significantly increased the average number of households per PSU in their respective Census divisions.

For variance estimation purposes, each self-representing PSU has, within its Census divisions, the approximate effect of two non-self-representing PSU's. Moreover, because of their large populations, the New York City, Chicago, and Los Angeles PSU's have the effect of three or four non-self-representing PSU's.

In general, the convention of treating the very large MSA's as self-representing PSU's, and the restriction treating Alaska and Hawaii as separate strata, means that the 116 strata used in the 1993 sample design were, for variance estimation purposes, more like a design with approximately 150 strata.[8]

Table 3a. Number of Primary Stage Strata and Population by Census Division for 1980, 1984, and 1993 RECS Sample Designs

Census Division	Number of Primary Stage Strata			Population (millions)		
	1980 Design	1984 Design	1993 Design	1980 Estimates	1980 Census	1990 Census
Total	131	129	116	218.1	226.5	248.7
New England	10	10	8	12.3	12.3	13.2
Middle Atlantic	14	15	13	36.8	36.8	37.6
East North Central	16	17	17	41.2	41.7	42.0
West North Central	17	17	9	17.0	17.2	17.7
South Atlantic	17	19	20	34.6	37.0	43.6
East South Central	13	12	8	14.0	14.7	15.2
West South Central	12	12	14	22.0	23.7	26.7
Mountain	13	11	10	10.3	11.4	13.7
Pacific	19	16	17	29.8	31.8	39.1

Table 3b. Average Population and Number of Households per Stratum by Census Division for 1980, 1984, and 1993 RECS Sample Designs

Census Division	Average Population per Primary Stage Stratum (millions)			Number of Households for 1993 Design (1990 Census Data) (millions)	
	1980 Design	1984 Design	1993 Design	Total	Average per Stratum
Total	1.7	1.8	2.1	91.9	0.8
New England	1.2	1.2	1.7	4.9	0.6
Middle Atlantic	2.6	2.5	2.9	13.9	1.1
East North Central	2.6	2.5	2.5	15.6	0.9
West North Central	1.0	1.0	2.0	6.7	0.7
South Atlantic	2.0	1.9	2.2	16.5	0.8
East South Central	1.1	1.2	1.9	5.7	0.7
West South Central	1.8	2.0	1.9	9.7	0.7
Mountain	0.8	1.0	1.4	5.0	0.5
Pacific	1.6	2.0	2.3	13.9	0.8

Source: *Residential Energy Consumption Survey Primary Stage Sample Design Plan (Third Draft)*, Sampling Section, Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), February 1, 1993.

[8] *Residential Energy Consumption Survey Second Stage Sample Design Plan*, Sampling Section, Survey Research Center, Institute for Social Research, page 7, University of Michigan (Ann Arbor, Michigan), June 1, 1993.

Note that the number of strata for the West North Central Census Division dropped from 17 in the 1980 and 1984 designs, to 9 in the 1993 design. The large number of strata (relative to its population), in this Census division in the 1980 design, was a result of the stipulation preventing the strata from crossing Census divisions or Federal region boundaries, as well as the need for the statistical reliability of data for each of these geographic domains.

The preliminary finding (based on energy data collected prior to 1978) showing the variance of household energy expenditures to be higher in this Census division than in other Census divisions, also contributed to this large number. The relatively high number of PSU's assigned to this division in the 1980 design was carried over to the 1984 design. In the 1993 design, the decision on the number of PSU's to assign to each division was made independently of previous designs. A study of the 1990 RECS results indicated that there was no significant difference in the variance of household energy expenditures for the different Census divisions. Consequently, the number of PSU's assigned to the West North Central Census Division was reduced in the 1993 design.

If the RECS sample design was optimized to produce the lowest possible variance for the estimated energy consumption per household in the entire Nation, and if the optimization analysis was developed under the assumptions that: (1) the sample size was fixed, (2) the standard deviation of household energy consumption was the same for all Census divisions, and (3) costs was the same for all Census divisions, then the number of strata in a Census division should be roughly proportional to its number of households. Inequalities occurred for the following reasons: (1) upper boundaries were placed on the variances for the per household estimates of energy expenditures in each of the Census divisions; and, (2) the convention of treating large MSA's as single PSU's was used. The upper boundaries resulted in the use of extra PSU's for those Census divisions with the smallest population. The convention of treating very large metropolitan areas as self-representing PSU's meant that their population was much larger than that of other PSU's.

Table 3 also shows that the average population per stratum for the nine Census divisions is closer in the 1993 sample design than in the 1980 or 1984 sample designs. This improvement followed the elimination of the ten Federal region estimates, as well as the assumption that the variance of household energy expenditures is approximately equal for all Census divisions.

Criteria for Forming Primary Sampling Units

The RECS sample designs all used the same overall strategy for defining PSU's. In metropolitan areas, the PSU's are usually defined by the boundaries of MSA's. In non-metropolitan areas, the PSU's are usually individual counties or groups of contiguous counties. This is the same convention that is used for many national surveys conducted by the Federal government.

In the 1980 and 1984 sample designs, PSU boundaries did not cross State lines, with the exception of the Washington, DC metropolitan area. The design included this restriction in order to provide for its possible future expansion to a State-level survey. In the 1993 sample design, this restriction was relaxed to allow PSU's to cross State lines in more situations. These exceptions always involved PSU's in those metropolitan areas that included counties in two or more States.

In order to provide for the contingency that EIA forecasting models may require energy consumption estimates for the larger States, PSU boundaries were *not* allowed to cross State lines in California, New York, Texas, and Florida.

In all RECS sample designs, the restriction that prevented PSU boundaries from crossing Census division lines was observed. This restriction reflected the desire to control the effect on the variance in using the RECS data to make Census division-level estimates. Moreover, in the 1980 design, PSU boundaries did not cross Federal region boundaries.

In the 1980 sample design, the 3,141 counties and independent cities in the Unites States were formed into 1,782 PSU's.[9] For the 1984 sample design, 1,799 PSU's were formed. For the 1993 sample design, 1,786 PSU's were formed. The following areas indicate where the PSU definitions were affected:

- Changes in the definitions of MSA's
- Changes in the treatment of MSA's that cross State lines
- Changes in the number of cases where large contiguous MSA's are combined to form a single PSU.

Criteria for Grouping Primary Sampling Units Into Strata

The PSU's are grouped into strata. Within the Census divisions, there are four types of strata:

1. ***Self-Representing Strata***: Each of these strata contains only one PSU, and is the geographic area corresponding to a MSA that has a large population. These PSU's are designated as self-representing or "certainty" PSU's because they are included in the sample with certainty.

2. ***Non-Self-Representing MSA Strata***: Each of these strata contains two or more PSU's. Each PSU contained in these strata is defined as the geographic area covered by a MSA with a relatively small or medium-size population.

3. ***Non-Self-Representing Non-MSA Strata***: Each stratum in this category contains two or more PSU's. Each PSU in these strata is comprised of a single non-metropolitan county (or similar jurisdiction), or a group of contiguous non-metropolitan counties.

4. ***Non-Self-Representing Mixed Strata***: Each stratum in this category contains two or more PSU's. One or more of the PSU's is defined as the geographic area covered by a MSA with a relatively small or medium-size population. One or more of the PSU's is comprised of a single non-metropolitan county (or similar jurisdiction), or a group of contiguous non-metropolitan counties.

Self-Representing Strata

Large MSA's are designated as self-representing or "certainty" PSU's because their MOS constitutes a substantial fraction of the MOS for their Census Division or "intersection" in the 1980 design. The minimum size for self-representing PSU's varies by Census division or by intersection.

The set of self-representing PSU's has changed with each major revision in the RECS sample design. In addition, the definition of individual self-representing PSU's has changed in relation to changes in MSA definitions. The 1980 design used the 1980 MSA definitions, the 1984 design used the June 1983 MSA definitions, and the 1993 design used the June 1990 MSA definitions. Figure 4 lists the self-representing PSU's used for the 1980, 1984, and 1993 RECS sample designs.

[9] Residential Energy Consumption Survey, *Sample Design and Methodology National Household Surveys: 1980, 1981, and 1982*, p 7, Response Analysis Corporation (Princeton, New Jersey), August 1983

Figure 4. Self-Representing PSU's for 1980, 1984, and 1993 RECS Sample Designs by Census Division

Census Division	Self-Representing PSU's		
	1980 Design	1984 Design	1993 Design
New England	Boston	Boston	Boston
	Hartford	Hartford	Hartford
	Providence	Providence	
	Fairfield County, CT	Fairfield County, CT	
	New Haven	New Haven	
Middle Atlantic	New York	New York	New York
	Philadelphia	Philadelphia	Philadelphia
	Nassau-Suffolk (Long Island)	Nassau-Suffolk (Long Island)	Nassau-Suffolk (Long Island)
	Pittsburgh	Pittsburgh	Pittsburgh
	Newark	Newark	Newark
East North Central	Chicago	Chicago	Chicago
	Detroit	Detroit	Detroit
	Cleveland	Cleveland	Cleveland
West North Central	Minneapolis	Minneapolis	Minneapolis
	St Louis	St Louis	St Louis
	Kansas City	Kansas City	Kansas City
South Atlantic	Washington, D.C.	Washington, D.C.	Washington, D.C.
	Atlanta	Atlanta	Atlanta
	Baltimore	Baltimore	Baltimore
		Tampa	Tampa
		Miami	Miami
East South Central	None	None	None
West South Central	Houston	Houston	Houston
	Dallas	Dallas	Dallas
Mountain	Phoenix	Phoenix	Phoenix
	Denver	Denver	Denver
	Salt Lake City	Salt Lake City	
Pacific	Los Angeles	Los Angeles	Los Angeles
	San Diego	San Diego	San Diego
	Orange County, CA	Orange County, CA	Orange County, CA
	Seattle	Seattle	Seattle
		Riverside-San Bernardino	Riverside-San Bernardino
	San Francisco-Oakland	San Francisco-Oakland	San Francisco Oakland
	Portland		Sacramento
	Tacoma		San Jose

Sources: • Energy Information Administration, Residential Energy Consumption Survey *Primary Stage Sample Design Plan* (Third Draft), Sampling Section, Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), February 1, 1993. • The 1984 Residential Energy Consumption Survey *Sample Design Procedures Manual*, Orkand Corporation (Silver Spring, MD), March 1986. • The Residential Energy Consumption Survey, *Sample Design and Methodology 1980 - 1981 National Household Survey* (Draft Report), Response Analysis Corporation (Princeton, NJ), June 1981.

Non-Self-Representing MSA Strata

For each RECS design, the non-self-representing MSA PSU's were grouped into strata. Where possible, each stratum consisted of MSA's of similar size, from the same State (or at least contiguous States), and with a common dominate main space-heating fuel. Examples of non-self-representing strata in the 1993 RECS sample design are as follows:

Example 1: (Medium-Size MSA's in New York)
- Buffalo
- Rochester
- Syracuse

Example 2: (Smaller MSA's in the northern half of the West North Central Census Division)
- Des Moines, Iowa
- Duluth, (Minnesota portion)
- Cedar Rapids, Iowa
- St Cloud, Minnesota
- Fargo-Moorhead (Both North Dakota and Minnesota portions)
- Davenport-Rock Island-Moline (Iowa portion)
- Waterloo-Cedar Falls, Iowa
- Sioux Falls, South Dakota
- Rochester, Minnesota
- Iowa City, Iowa
- Bismarck, North Dakota
- Dubuque, Iowa
- Rapid City, South Dakota
- Grand Forks, North Dakota

In Example 2, the Wisconsin portion of the Duluth MSA is not in the West North Central Census Division; hence, it was placed in a non-self-representing MSA stratum in the East North Central Census Division. Similarly, the Illinois portion of the Davenport-Rock Island-Moline MSA was placed in a non-self-representing MSA stratum in the East North Central Census Division. However, the entire Fargo-Moorhead MSA is in the West North Central Census Division; hence, this PSU was allowed to cross State lines.

Non-Self-Representing Non-MSA Strata

The non-self-representing, non-MSA PSU's were also grouped into strata. The strata for these PSU's in the 1980 and 1984 sample designs were based heavily on the strata for the Current Population Survey, with some adjustments made for climate. As a result, the PSU's were grouped according to socioeconomic more than energy-related variables. However, the more important factor observed in devising each design was the need to have strata of smallest variation in aggregate measure of size for each Census division or each 'intersection' in the 1980 design. For example, the Monroe County, Florida PSU (The Everglades National Park and the Florida Keys) was placed in the same stratum as the PSU for Cherokee, Clay, Graham, and Swain counties in North Carolina (The Great Smoky Mountains National Park and surrounding areas). In the 1993 design, the PSU's were grouped into strata using energy-related variables. In particular, climate and main space-heating fuel where used in forming the strata.

Non-Self-Representing Mixed Strata

All RECS sample designs used some primary stage strata that contained metropolitan PSU's as well as non-metropolitan PSU's. These strata are called mixed strata. The use of mixed strata does cause problems in the

weighting and variance estimation procedures. For this reason, the number of mixed strata was kept at a minimum.

Because of the unique climates of Hawaii and Alaska, all RECS sample designs were constructed such that one PSU from Alaska and one PSU from Hawaii would be selected during the primary stage. This was accomplished by forming a mixed stratum composed only of PSU's in Hawaii, both metropolitan and non-metropolitan, and selecting a PSU from the stratum. Similarly, another PSU was selected from a mixed stratum composed only of PSU's in Alaska, both metropolitan and non-metropolitan.

In the 1980 RECS sample design, there were four additional mixed strata. These resulted from the mandate preventing strata boundaries from crossing any of the 17 intersections defined earlier. The four mixed strata were formed using (1) all PSU's in New Mexico, (2) all PSU's in Idaho, (3) all PSU's in Minnesota (with the exception of the Minneapolis-St Paul MSA), and (4) all MSA PSU's in North Dakota and South Dakota plus, some non-MSA PSU's in these two States.

Number of Strata by Type and Census Division

The total number of strata for the RECS sample design reflects decisions made in the following areas:

- The optimization analysis
- The number of large metropolitan areas that are designated as certainty PSU's
- The minimum number of PSU's for the Census divisions
- How well noncertainty PSU's can be grouped into strata of similar type, size, and energy characterization.

An additional feature of the 1993 design was the stipulation requiring, for each Census division, the use of an even number of strata for both the set of noncertainty metropolitan PSU's, and the set of noncertainty nonmetropolitan PSU's. The only exception to this was noncertainty metropolitan PSU's in the South Atlantic Census Division. This exception was due to budgetary restrictions.

The reason for adding this feature to the 1993 RECS sample design was to improve the variance estimation procedures. In particular, the variance procedure used for RECS, the Balanced Repeated Replication (BRR), requires the noncertainty strata to be paired and treated, for variance estimation purposes, as if the PSU's selected from the two strata were a sample size of two from a single large stratum.[10] Using an even number of strata for each noncertainty type facilitates their pairing.

In the 1980, 1981, and 1982 RECS, a total of 131 primary stage strata were used. Since the sample design for the 1984 RECS combined the 1980 and 1984 designs, it contained 153 PSU's. This breaks down to 107 PSU's that were used in both the 1980 and 1984 designs, 24 PSU's that were used in the 1980 design and the 1984 data collection, but not in the 1984 design, and 22 PSU's that were used only in the 1984 design. The 1987 and 1990 RECS each used a total of 129 strata. In the 1993 RECS, a total of 116 strata were used.

As mentioned earlier, a double strata were created in the West South Central Census Division in the 1993 RECS primary stage design. The double strata consisted of the Fort Worth, Texas MSA; the San Antonio,

[10] Kalton, G., "Practical Methods for Estimating Survey Sampling Errors," *Bulletin of the International Statistical Institute:* 47,3: 495-524, 1977.

Texas MSA; and the Austin, Texas MSA. Two PSU's were selected from these strata. The double strata were created in this situation for the following reasons:

- The stipulation preventing the primary stage strata from crossing Texas State boundaries
- The desire to have an even number of non-self-representing MSA strata in each Census division
- The desire to place non-self-representing MSA's into strata with MSA's of approximately the same size
- The desire to use fewer strata for the 1993 RECS design than were used for the 1984 RECS design.

One alternative would have been to define the Fort Worth MSA, the San Antonio MSA, and the Austin MSA each as a self-representing PSU. This would have increased the number of strata for the 1993 RECS design by one. Another alternative would have been to pair each of these three large MSA's with one or more smaller MSA's, to form non-self-representing strata; but this would run counter to the goal of forming strata with PSU's of roughly comparable size.

Tables 4, 5, and 6 give the number of primary stage strata by Census division and type of PSU, for the 1980, 1984, and 1993 designs. Appendix C contains the strata listings by Census Division for the 1980, 1984 and 1993 sample designs.

Table 4. Number of Strata for PSU's by Census Division and Type of Strata for the 1980 Design

Census Division	Total	Self Representing	Non-Self Representing (NSR)		
			MSA	Non-MSA	Mixed
Total	131	31	53	41	6
New England	10	5	3	2	0
Middle Atlantic	14	5	7	2	0
East North Central	16	3	9	4	0
West North Central	17	3	5	7	2
South Atlantic	17	3	8	6	0
East South Central	13	0	6	7	0
West South Central	12	2	6	4	0
Mountain	13	3	3	5	2
Pacific	19	7	6	4	2

Source: Energy Information Administration, Residential Energy Consumption Survey, "Sample Design and Methodology 1980 - 1981 National Household Survey" (Draft Report), Response Analysis Corporation (Princeton, NJ), June 1981.

Table 5. Number of Strata for PSU's by Census Division and Type of Strata for the 1984 Design

Census Division	Total	Self Representing	Non-Self Representing (NSR)		
			MSA	Non-MSA	Mixed
Total	129	32	54	41	2
New England	10	5	3	2	0
Middle Atlantic	15	5	8	2	0
East North Central	17	3	9	5	0
West North Central	17	3	5	9	0
South Atlantic	19	5	8	6	0
East South Central	12	0	6	6	0
West South Central	12	2	6	4	0
Mountain	11	3	3	5	0
Pacific	16	6	6	2	2

Source: Energy Information Administration, "1984 Residential Energy Consumption Survey Sample Design Procedures Manual," The Orkand Corporation (Silver Spring, MD), March 1986.

Table 6. Number of Strata for PSU's by Census Division and Type of Strata for the 1993 Design

Census Division	Total	Self Representing	Non-Self Representing (NSR)		
			MSA	Non-MSA	Mixed
Total	116	31	51	32	2
New England	8	2	4	2	0
Middle Atlantic	13	5	6	2	0
East North Central	17	3	10	4	0
West North Central	9	3	2	4	0
South Atlantic	20	5	9	6	0
East South Central	8	0	4	4	0
West South Central	14	2	8	4	0
Mountain	10	2	4	4	0
Pacific	17	9	4	2	2

Source: Energy Information Administration, 1993 Residential Energy Consumption Survey, *Primary Stage Sample Design Plan* (Final Draft), Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), June 1993.

Primary Sampling Unit Selection Procedures

For each RECS sample design, a single PSU in each non-self representing stratum was selected using probability-proportional-to-size (PPS) sampling; a procedure used in many other surveys conducted by the Federal Government. In PPS sampling, each PSU was assigned a MOS. In the 1980 design, the MOS of a PSU was equal to 1978 PSU population estimates [CPS, Series P-26, No.783, issued February 1980], in the 1984 design, the MOS was equal to the 1980 Census population counts, and in the 1993 design, the MOS was equal to the 1990 Census count of occupied housing units. Because there is a one-to-one correspondence between occupied housing units and households, the number of occupied housing units in a PSU is equal to the number of households.

However, this change from population to number of households for the MOS of a PSU in the 1993 design, was instituted because the elementary-sampling unit for all RECS sample designs was a housing unit and not

a person. While RECS has always used the housing unit as its elementary-sampling unit, previous designs used population as the MOS for PSU's. There is a strong relationship between the population of a PSU and the number of occupied housing units in a PSU. As a result, this change has had only a minor effect.

The probability of selecting a particular PSU is proportional to its MOS, more specifically, its MOS divided by the MOS total for all PSU's in the stratum.

In the 1980 design, the PSU's were newly selected. Controlled selection was used to increase the assurance of a reasonably close match between the distribution of the sample, and actual population by State.[11]

As a cost-saving feature in the 1984 design, a Keyfitz-type procedure, developed at the Bureau of the Census,[12] was used to ensure that a large percentage of the PSU's selected for the 1980 design would be carried over to the 1984 design. In fact, of the 129 PSU's selected for the 1984 design, 107 of these had been used for the 1980 design, leaving only 22 PSU's which were not used in the 1980 design. In particular, the 97 non-self-representing PSU's selected for the 1984 design were chosen in such a way that 76 of these had also been selected for the 1980 design, with 21 having been selected for the first time. Of the self-representing PSU's selected for the 1984 design, only one was not used in the 1980 design. The remaining self-representing PSU's were used in the 1980 design either as self-representing or non-self-representing PSU's. Controlled selection was also used in the 1984 design to ensure that the PSU's would be distributed among the States.

The original plans for the 1993 RECS sample design called for the selection of PSU's to be independent of the 1980 and 1984 selections. However, due to budget reductions, the plans were changed and again a Keyfitz-type procedure[13] was used to ensure that a large proportion of the PSU's chosen for the 1984 design were carried over to the 1993 design. In fact, of the 116 PSU's selected for the 1993 design, 94 of these were used in the 1984 design, with only 22 having been selected for the first time.

Of the 116 strata used for the 1993 design, 31 contained single self-representing PSU's, while 85 contained multiple PSU's. Of the 31 self-representing PSU's in this design, 27 were self-representing PSU's, and 2 were non-self-representing PSU's in the 1984 design. The remaining 2 self-representing PSU's in the 1993 design (San Francisco and Oakland) were used in the 1984 design as a single self-representing PSU. The 85 non-self-representing PSU's selected for the 1993 design were chosen in such a way that 63 of these had also been selected for the 1984 design; the other 22 non-self-representing PSU's were first time selections.

[11]Goodman, Roe, and Kish, Leslie. "Controlled Selection: A Technique in Probability Sampling." *Journal of the American Statistical Association* Vol. 45, September 1950, pp. 350-372.

[12]1970 *Current Population Survey Redesign: Illustration of Computation of PSU Probabilities Within a Given 1970 Stratum*, Memorandum from W.M. Perkins to J. Waksberg, February 19, 1971. • The *1970 Current Population Survey Redesign: Proposed Method for Deriving Sample PSU Selection Probabilities Within 1970 Non-Self-Representing Strata*, Memorandum from W.M. Perkins to J. Waksberg, August 5, 1970.

[13]Kish, Leslie, and Scott, Alastair, "Retaining Units After Changing Strata and Probabilities." *Journal of the American Statistical Association*, Vol. 66, Number 335, Applications Section, September 1971.

Summary of Primary Stage

For each RECS design, Table 7 summarizes the universe, the number of PSU's, the MOS used in selecting the PSU's, and the basis for the MSA definitions.

Table 7. Comparison of RECS Sample Designs at the PSU Level

Survey Design Year	Universe	Number of Strata	*Measure-of-Size* for PSU's	*Definition of* MSA's
NIECS (1978) and Screener (1979)	48 contiguous States and District of Columbia	103	1970 Census	1970 Census definitions
1980	50 States and District of Columbia	131	1978 population estimates from the Bureau of the Census	1980 definitions
1984	50 States and District of Columbia	129 total 107 retained from 1980 RECS Design 22 selected for the first time for 1984 design	Population: 1980 Census	June 1983 definitions
1993	50 States and District of Columbia	116 total 94 retained from 1984 design 22 selected for the first time for 1993 design	Number of households: 1990 Census	June 1990 definitions

Sources: Energy Information Administration, Residential Energy Consumption Survey, "Sample Design and Methodology 1980 - 1981 National Household Survey" (Draft Report), Response Analysis Corporation (Princeton, NJ), June 1981. • The National Interim Energy Consumption Survey, Part I: Methodology on Household and Utility Company Surveys, Response Analysis Corporation (Princeton, NJ), June 30, 1981. • The 1984 Residential Energy Consumption Survey *Sample Design Procedures Manual*, Orkand Corporation (Silver Spring, MD), March 1986. • The 1993 Residential Energy Consumption Survey, *Primary Stage Sample Design Plan*" (Final Draft), Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), June 1993.

6. Secondary Stage of the RECS Sample Design

This chapter discusses the secondary stage of the RECS sample design. In this stage, each PSU that was selected during the primary stage is divided into SSU's, a sample of SSU's is selected from the PSU's, the selected SSU's are divided into listing segments, and a single listing segment is selected from each SSU. As part of the secondary stage the following decisions are made:

- How to integrate the new construction supplement into the secondary stage
- The number of SSU's
- Which criteria are to be used to form the SSU's
- How to select the SSU's?

Each of these decisions will be discussed in detail in the following sections.

1993 New Construction Supplement

The 1993 RECS sample design includes a supplemental sample of SSU's that contain a high proportion of new housing units. For sampling purposes, a new housing unit is defined as units that were first occupied within 6 years of the data collection period. From the start of construction to occupancy is typically 6 to 9 months; hence, for the 1993 RECS, housing units will be defined as new housing units if construction began on or after January 1, 1987. Because respondents may report the date that the housing units was first occupied instead of the date construction began, new housing units should closely correspond to those where the respondent classified the units as being constructed during the January 1988 to December 1993 period. The object of including a new construction supplement in the 1993 RECS sample design is to increase the number of interviews for new housing units.

For the 1990 RECS, new housing units were defined as those units where the respondent reported that the housing units were constructed during the January 1985 to December 1990 period. Because the sampling was finalized during the summer of 1990 and most interviews were conducted during the fall of 1990, the units that were first occupied during the latter part of 1990 would be underrepresented in the 1990 RECS.

The 1990 RECS revealed that there were approximately 400 new housing units in its sample of 5,095 housing units. The 1993 core sample of 5,000 housing units is also expected to have approximately 400 new housing units. The new construction supplement is designed to increase the sample of new housing units to approximately 1,200 units.

This supplement will address new construction that is interspersed throughout existing residential areas, as well as very recent, large-scale construction projects (i.e., subdivisions, apartment complexes, mobile home parks, etc.). The approach EIA has taken is one that combines an oversampling of new housing units in the core 1993 SSU's, with a supplemental sampling of SSU's from geographic areas showing a high density of new housing units.

Based on experience with the 1990 RECS, the core sample of 1,460 SSU's is expected to include about 200 SSU's that are located in areas with a high density of new housing units. The supplemental sample of 150 new construction SSU's will bring the total number of 1993 SSU's having significant new construction to about 350.

To reduce SSU listing costs, EIA used existing SSU listings in 30 of the 116 PSU's. In addition, the complete sample of 1,610 new SSU selections would not be fully implemented until just prior to the 1996 data collection. Consequently, a total of 149 supplemental new construction SSU selections were listed for the 1993 sample. Of this total, 105 were supplemental new housing unit selections from PSU's where all SSU's, core and new construction, were newly listed. The remaining 44 SSU's were selected from 30 PSU's where existing SSU listings were used. The complete 1993 RECS sample includes 348 SSU's from areas with a high density of new housing units.

Number of Secondary Sampling Units

The total number of PSU's and SSU's is determined by an optimization analysis, by budget constraints, and by other factors; determinations are usually made prior to a major revision in the sample design. Also, these decisions on the number of PSU's, the average number of SSU's per PSU, and the average number of housing units in an ultimate cluster, usually occur simultaneously.

While the total number of SSU's used in the RECS sample design has remained relatively constant, the number of SSU's assigned to some Census divisions has changed. Table 8 is divided into two parts (a and b), which lists, by each Census division, the population, the number of SSU's, and the average population per SSU. These data are given for the 1980, 1984, and 1993 designs. For the 1993 design, Table 8 also lists the number of households, and the average number of households per SSU for each Census division.

While Table 3 deals with PSU's and Table 8 deals with SSU's, they show similar trends. For example, the 1993 sample design resulted in fewer PSU's and SSU's in the West North Central Census Division. Also, the differences in the average population per SSU between the nine Census divisions, are smaller in the 1993 sample design than in earlier designs.

The average number of SSU's per PSU has remained relatively constant from the 1980 sample design through the 1984 and 1993 sample designs; this especially applies to those SSU's in the core sample. Table 9 lists the number of core SSU's, the number of PSU's, and the average number of core SSU's per PSU by Census division, for each major revision of the RECS sample design.

Once the number of SSU's for a Census division has been determined, they are assigned to PSU's in a manner that is approximately proportionate to the MOS for the corresponding stratum. For example, if Stratum A contains 12 percent of the households in the Census division, then approximately 12 percent of the SSU's assigned to that Census division are assigned to the PSU that is selected from Stratum A. In the 1993 design, the number of SSU's for each stratum was rounded to an even number to make variance calculations easier.

Table 8a. Number of Secondary Sampling Units and Population by Census Division for 1980, 1984, and 1993 RECS Sample Designs

Census Division	Number of Secondary Sampling Units			Population (millions)		
	1980 Design[a]	1984 Design	1993 Design[a]	1980 Estimates	1980 Census	1990 Census
Total	1,515	1,516	1,460	218.1	226.5	248.7
New England	128	122	116	12.3	12.3	13.2
Middle Atlantic	191	202	202	36.8	36.8	37.6
East North Central	195	202	202	41.2	41.7	42.0
West North Central	172	202	118	17.0	17.2	17.7
South Atlantic	178	202	204	34.6	37.0	43.6
East South Central	128	122	118	14.0	14.7	15.2
West South Central	128	122	152	22.0	23.7	26.7
Mountain	153	122	126	10.3	11.4	13.7
Pacific	242	220	222	29.8	31.8	39.2

Table 8b. Average Population per Secondary Sampling Units and Number of Households by Census Division for 1980, 1984, and 1993 RECS Sample Designs

Census Division	Average Population per Secondary Sampling Unit (millions)			Number of Households for 1993 Design[a] (1990 Census Data) (millions)	
	1980 Design[a]	1984 Design	1993 Design[a]	Total	Average per SSU
Total	0.14	0.15	0.17	91.9	0.064
New England	0.10	0.10	0.11	4.9	0.043
Middle Atlantic	0.19	0.18	0.19	13.9	0.069
East North Central	0.21	0.21	0.21	15.6	0.077
West North Central	0.10	0.09	0.15	6.7	0.057
South Atlantic	0.19	0.18	0.21	16.5	0.081
East South Central	0.11	0.12	0.13	5.7	0.047
West South Central	0.17	0.19	0.18	9.7	0.064
Mountain	0.07	0.09	0.11	5.0	0.040
Pacific	0.12	0.14	0.16	13.9	0.063

[a]Core sample only.

Sources: •Energy Information Administration, Residential Energy Consumption Survey, *Sample Design and Methodology 1980-1981 National Household Survey* (Draft Report), Response Analysis Corporation (Princeton, NJ), June 1981. • The 1993 Residential Energy Consumption survey, *Second Stage Sample Design Plan* (Final Draft), Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), June 1993.

Table 9. Number of Secondary Sampling Units, Primary Sampling Units, and SSU's per PSU by Census Division for 1980, 1984, and 1993 RECS Sample Designs

Census Division	Number of Secondary Sampling Units			Number of Primary Sampling Units			Number of Secondary Sampling Units per Primary Sampling Unit		
	1980 Design[a]	1984 Design	1993 Design[a]	1980 Design	1984 Design	1993 Design	1980 Design[a]	1984 Design	1993 Design[a]
Total	1,515	1,516	1,460	131	129	116	11.6	11.8	12.6
New England	128	122	116	10	10	8	12.8	12.2	14.5
Middle Atlantic	191	202	202	14	15	13	13.6	13.5	15.5
East North Central	195	202	202	16	17	17	12.2	11.9	11.9
West North Central	172	202	118	17	17	9	10.1	11.9	13.1
South Atlantic	178	202	204	17	19	20	10.5	10.6	10.2
East South Central	128	122	118	13	12	8	9.8	10.2	14.7
West South Central	128	122	152	12	12	14	10.7	10.2	10.9
Mountain	153	122	126	13	11	10	11.8	11.1	12.6
Pacific	242	220	222	19	16	19	12.7	13.8	11.7

[a] Core sample only.

Sources: •Energy Information Administration, Residential Energy Consumption Survey, *Sample Design and Methodology 1980 - 1981 National Household Survey* (Draft Report), Response Analysis Corporation (Princeton, NJ), June 1981. • *The 1984 Residential Energy Consumption Survey Sample Design Procedures Manual*, The Orkand Corporation (Silver Spring, MD), March 1986. • The 1993 Residential Energy Consumption Survey, *Primary Stage Sample Design Plan* (Final Draft), Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), June 1993. • The 1993 Residential Energy Consumption Survey, *Second Stage Sample Design Plan* (Final Draft), Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor), Michigan, June 1993.

Criteria for Forming Secondary Sampling Units

For each RECS design, the SSU's are well-defined geographic areas whose boundaries correspond to the boundaries of small geographic units that were defined by the Bureau of the Census. The 1980 and 1984 sample designs used Census tracts, enumeration districts, individual blocks and block groups. In the 1993 RECS sample design, the SSU's are usually Census blocks or a set of contiguous Census blocks. A Census block is "An area bounded on all sides by visible features such as streets, roads, streams, and railroad tracks, and occasionally by nonvisible boundaries such as city, town, or county limits, property lines, and short imaginary extensions of streets. Blocks do not cross Census tract or block numbering area boundaries. A block is the smallest geographic tabulation area from the 1990 Census."[14]

There are several major advantages for using Census units to define SSU's:

- Census data can be used to define a MOS for the SSU's
- Small area Census units exist for the entire country
- Census unit boundaries are well-defined and easy to locate

[14] U.S. Department of Commerce. Bureau of the Census, *1990 Census of Population and Housing Tabulation and Publication Program*, July 1989, p. 49.

- Maps of the Census units can be obtained from the Bureau of the Census, or created, using the Topologically Integrated Geographic Encoding and Referencing (TIGER) System data files obtained from the Bureau of the Census.

For the first time in the 1990 Census, the Bureau of the Census defined blocks for the entire country. In the 1980 Census, the smallest Census unit in many sparsely populated parts of the country was an enumeration district. (Census tracts and blocks were not defined for sparsely populated rural areas.) Consequently, some rural SSU's were enumeration districts in the 1980 and 1984 sample designs.

Number of Housing Units in SSU's

The target number of households in a SSU was different for the 1993 RECS design than for the 1980 and 1984 RECS designs. For the 1993 RECS, the core SSU's were defined so that their projected minimum number of housing units was 50, and the New Construction Supplement SSU's were defined so that their projected minimum number of housing units was 96. For the 1980 and 1984 RECS sample designs, the SSU's were defined so that the target number of housing units was 400.

The 1980 and 1984 RECS used the larger minimum number of housing units in a SSU, in order to facilitate the new construction update procedures used in RECS cycles between major redesign efforts. For the 1996 and subsequent RECS, the new construction update procedures will cover not only the selected SSU's, but may also cover neighboring SSU's.

The smaller size of the 1993 RECS sample design was used to reduce the number of SSU's in which rough-counting procedures, segment formulation procedures, and segment selection procedures were necessary. In fact, for the majority of the SSU's selected in the 1993 design, these three sets of procedures were not necessary. Consequently, the majority of the SSU's in this design were equal to their listing segment.

Secondary Sampling Unit Selection Procedures

The following steps summarize the procedures used in selecting the SSU's:

- Assigning a MOS to the SSU's
- Stratify or order the SSU's prior to selection
- Randomly selecting the SSU's.

MOS for SSU's

During the secondary stage, the SSU's are selected using probability-proportionate-to-size sampling (PPS). In order to use PPS sampling, each SSU is assigned a MOS. The MOS used for the NIECS and Screener was the 1970 Census population of the SSU. Since the 1980 Census data were not available, the MOS used for the 1980 RECS sample design was the estimated number of households per SSU. These estimates were derived by combining information from a number of sources: The 1970 Census results, National Planning Data Corporation, Reuben H. Donnelly Company's marketing data on the number of households, and contacts with local officials.

The SSU's in the core sample for the 1980 design, along with their selection probabilities, were also used for the 1981 and 1982 RECS. Similarly, those SSU's that were carried over from the 1980 design to the 1984 design also carried over with them their original selection probabilities. This differs from the Keyfitz

procedure used for selecting PSU's; a procedure which usually altered selection probabilities even when original PSU's were reselected. In those PSU's that were retained, the SSU stage, including its probability of selection, was frequently carried over intact from the earlier design. In those PSU's that were either altered to reflect changes in their definition, or were affected when the SSU structure was totally revised, the MOS used to select the SSU was based on the 1980 Census. In the 1993 RECS redesign effort, the 1990 Census number of housing units was used to obtain the MOS for SSU's.

SSU Stratification or Ordering Procedures

In all three designs, the SSU's were ordered or stratified prior to selection. The 1980 and 1984 designs used the intermediate step of selecting a Minor Civil Division as an mechanism for guaranteeing that SSU's are distributed across the PSU's. The 1993 RECS design also stratified the SSU's prior to selection. In non-MSA PSU's, the SSU's were stratified by geographic location. In MSA PSU's, the SSU's were first stratified by predicted energy expenditures, and within these strata, by geographic location. The SSU's for non-MSA PSU's were not stratified by predicted energy expenditures, because in most of these areas there is too little housing unit clustering, by variables related to energy expenditures, to justify these procedures.

SSU Selection Procedures for the 1980 and 1984 RECS Sample Designs

The SSU's for the 1980 and 1984 designs were selected in a two-step procedure:

1. ***Minor Civil Divisions*** (MCD) such as cities, towns, and other Census units were selected from each RECS PSU

2. ***Secondary-Sampling Units*** (SSU's), which are Census tracts, block groups, and enumeration districts, were selected from each MCD chosen in the first step. Usually only one SSU was selected, but in some cases multiple SSU's were selected. The SSU's were selected using PPS sampling, in which Census counts or estimates of the number of households were used as a MOS.

The MCD's were selected by using a systematic sample, and were intended to stratify the SSU's into categories reflecting the size and characteristics of the MCD. For each PSU, this was accomplished in the following manner:

- Calculating the sum of the MOS for all MCD's in the PSU. (For notational purposes, denote the sum as $\sum$ MOS.)

- Setting the PSU "zone" interval equal to $\sum$ MOS, divided by the number of SSU's to be selected from the PSU. (For notational purposes, denote the "zone" interval as W.)

- Ordering the MCD's; first order the MCD by type, then by size (MOS) or by geographic location

- Using the ordered MCD's and their MOS to create "paper" zones for the PSU. Each "paper" zone contains W households. The first zone extended from the "first" household in the PSU to the W-th household. The second zone extended from household W + 1 to household 2W. Any MCD might lie entirely in a single zone, overlap a "paper" zone boundary and be partly included in two zones, or extend through all or part of three or more zones

- For each zone, select a random number between 1 and W. For zone i denote the random number as K_i. The MCD corresponding to the K_i-th household in the i-th zone is thereby selected.

The definitions for types of MCD's, and the procedures for ordering the MCD's within type classes, are given below:

- The MCD's for central cities are first. Within this group, MCD's are ordered by their MOS (largest to smallest)
- Other MCD's in urbanized areas follow. Within this group MCD's are ordered by geographic location
- Those MCD's, the majority of whose population lies in urban areas, form the next class. Within this group, the MCD's are ordered by their MOS (largest to smallest)
- All other MCD's follow. These MCD's are ordered by geographic location.

This ordering also guarantees that those selected represent the different types of MCD's.

Any MCD that is placed in two or more "paper" zones could be selected from each zone. This is most likely to happen to MCD's that have a large MOS, in particular, to MCD's that are central cities in MSA's.

SSU Selection Procedures for the 1993 RECS Sample Design

Beginning with the 1993 RECS sample design, a substantial change was made in the SSU selection procedures. A decision was made to stratify and/or order the SSU's according to their energy-related characteristics and geographic location. (The 1980 and 1984 designs used urban/rural status as well as geographic location.) Using predicted energy expenditures to stratify the SSU's prior to selection will increase the precision of energy expenditure estimates, as well other energy-related estimates.

Within the PSU's selected for MSA strata (including certainty PSU's), the SSU's are first stratified by predicted energy expenditures. Within the strata determined by predicted energy expenditures, the SSU's are ordered geographically by county, Census tract, or block group. Within the PSU's selected for non-MSA strata, the SSU's are ordered geographically, as they are not first stratified by energy-related characteristics.

Respondents' estimates of energy expenditures were collected in the 1990 Census long form, but this information was not used to stratify the SSU's. The Census energy expenditure data only covered expenses incurred by the household; it did not take into account expenses that were included in the rent. In order to make full use of the Census data on expenditures, it would be necessary to estimate the energy expenditures in those households where utilities are included as part of the rent. An additional problem arises in cases where the rent includes only part of the utilities. It would be difficult to accurately detect these cases using Census data, or to calculate the energy expenditures included in the rent. Finally, Census expenditure data are known to be upwardly biased.

The predicted energy expenditures were determined by an equation that was developed using data from the 1990 RECS. The independent variables used in the equation were available from both the 1990 RECS data base and the 1990 Census. Census data was used to determine the values for the predictor variables. This equation was applied at the Census tract level and, consequently, all SSU's (usually blocks) in the same tract were placed in the same stratum, having the same predicted average energy expenditures.

The energy expenditures model used the following four variables:

AVEHI = Average Household Income for Households in the Census Tract (in thousands of dollars)

AVENHM = Average Number of Household Members for Households in the Census Tract

PSFDHU = Proportion of Single-Family Detached, Housing Units among all Housing Units in the Census Tract

AVENR = Average Number of Rooms for Occupied Housing Units in the Census Tract.

A regression procedure was used to develop the following equation, which predicts a log of the energy expenditures (LNEXP) based on the above four variables:

$$\begin{aligned} LNEXP = {} & 6.1384 + (.09179 \times AVENR) \\ & + (.06765 \times AVENHM) \\ & + (.00678 \times PSFDHU) \\ & + (.00485 \times AVEHI) \end{aligned}$$

This equation was used to predict the average energy expenditures for each Census tract in the MSA-PSU's selected for the 1993 RECS redesign effort. The predicted energy expenditures that resulted were used to divide the SSU's into strata. These strata correspond to the upper, middle, and lower third percentiles for energy expenditures in each Census division.

The low-expenditure stratum contained those SSU's in Census tracts where the regression estimate of the average predicted energy expenditures for households fell below the 33.33rd percentile. The middle-expenditure stratum contained those SSU's in Census tracts where the estimate fell between the 33.33rd and 66.67th percentiles. The high-expenditure stratum contained those SSU's in Census tracts where the estimate fell above the 66.67th percentile.

Within the low, middle, and high-expenditure strata, the SSU's were ordered by geographic characteristics (county, tract, and block group). This ordering is used in the same way that stratification is used to distribute the selected SSU's geographically.

The geographic ordering of SSU's occurs first by county, within counties they are ordered by tract, and finally within tracts they are ordered by block. For MSA PSU's, the counties are ordered from the center of the MSA to the peripheries, keeping adjacent counties together. Most non-MSA PSU's consist of a single county, or very few counties. Hence, the ordering for non-MSA PSU's may be purely arbitrary. The tract and block numbering systems assigned by the Bureau of the Census orders the tracts geographically within a county, and orders the blocks geographically within the tracts.

A systematic selection procedure (similar to the procedures used to select SSU's for the 1980 design) is used to select the SSU's in the non-MSA PSU's and in each energy expenditure stratum in the MSA PSU's. In particular, the total MOS for each PSU or stratum is calculated (denote the total MOS by ΣMOS), the "zone" interval is calculated (denote the "zone" interval by W), and a random start is chosen. The total MOS is an estimate of the total number of households in the PSU or stratum. W equals ΣMOS divided by the number of SSU's that are to be selected from the PSU or stratum. Assume that for each SSU we have a list of households, these lists are combined to form a list for the PSU or stratum using the ordering of the SSU's. The PSU or stratum is divided into paper "zones" where the first "zone" contains the first W households in the list, the second "zone" contains the second W households in the list and, so forth. If K is the random start, then the SSU corresponding to the K-th household in each "zone" is selected. (For the 1980 design, a separate random value (K_i) was selected for each "zone." This was a minor change, which made processing easier.)

Additional Details of Secondary Stage for Each RECS

The 1980 RECS sample design used a total of 1,667 SSU's. 1,515 SSU's were in the core sample, and 152 SSU's were selected to be used in a reliability study. Initial planning called for the housing units selected from

these 152 SSU's to be interviewed twice for the 1980 RECS. This would have enabled EIA to analyze the reliability of the data collected for RECS. However, the plans for the second interview were eliminated for budgetary reasons. These 152 SSU's were used only for the 1980 RECS.

The sample design for the secondary stage to the 1981 RECS was the same as the sample design for the 1980 RECS, except that 152 SSU's originally incorporated into the 1980 RECS design for use in a reliability study, which were eliminated.

The secondary stage for the 1982 RECS was the same as that for the 1981 RECS.

As a result of the overlap in PSU's for the 1980 design and the 1984 RECS design, 1,249 of the 1,516 SSU's selected for the 1980 design were also used for the 1984 design. The SSU's in those PSU's that were selected for the first time in the 1984 sample design, were newly developed. However, these SSU's were not necessarily used in the 1984 RECS. Because a longitudinal design was incorporated in the RECS design (see Chapter 7), the SSU's assigned to the returning panel were not used until the 1987 RECS. The returning panel used SSU's from the 1982 RECS, even when those SSU's were in PSU's that were eliminated from the 1984 RECS design.

The secondary stage developed during the 1984 RECS design effort was used in its entirety for the first time in the 1987 RECS.

The secondary stage for the 1990 RECS is the same as that used for the 1987 RECS.

In some PSU's that were used in both the 1984 and the 1993 designs, the corresponding SSU's that were selected for the 1984 design were also used in the 1993 RECS. This was done for economic and expediential reasons. In fact, of the 1,460 SSU's in the core sample for the 1993 design, 281 were first selected for the 1980 design, 78 were first selected for the 1984 design, and 1,101 SSU's were initially selected for the 1993 design.

The SSU's that were selected during the 1993 design effort will be used for the 1996 RECS. All of the SSU's that were selected for the 1984 design will be eliminated.

Summary of Secondary Stage

Table 10 lists the number of SSU's and the MOS for the SSU's.

Table 10. Comparison of RECS Sample Designs at the SSU Level

Survey	Number of SSU's	Number of SSU's per PSU	Measure-of-Size for SSU's
NIECS (1978) and Screener (1979)	456	4.4	1970 Census population for 400 SSU's and amount of new construction for 56 SSU's
1980 RECS Sample Design	1,667 total:	Total sample: 12.7	Estimated number of households in 1980. Derived from a combination of 1970 Census data, National Planning Corporation estimates, and other sources.
	1,515 SSU's in core sample	Core sample: 11.6	
	152 SSU's in supplemental sample selected for reliability study		
1984 RECS Sample Design	1,516 total:	11.8	
	1,249 retained from 1980 design core sample		1984 SSU's retained from the 1980 core sample: estimated number of households in 1980 (same MOS used in 1980)
	267 selected for the first time for 1984 Design		SSU's selected for the first in 1984 estimated number of households in 1980. Derived from a combination of 1980 Census data, National Planning Data Corporation estimates, and other sources
1993 RECS Sample Design	1,610 Total:	Total sample:13.8	Number of housing units from 1990 Census
	1,460 for core sample	Core sample: 12.5	
	150 for New Construction Supplement		

Sources: •Energy Information Administration, Residential Energy Consumption Survey, *Sample Design and Methodology 1980 - 1981 National Household Survey* (Draft Report), Response Analysis Corporation (Princeton, NJ), June 1981. •The National Interim Energy Consumption Survey, Part I: Methodology on Household and Utility Company Surveys, Response Analysis Corporation (Princeton, NJ), June 30, 1981. • The 1993 Residential Energy Consumption survey, *Second Stage Sample Design Plan* (Final Draft), Survey Research Center, Institute for Social Research, University of Michigan (Ann Arbor, Michigan), June 1993.

7. Elementary-Unit Stage

During the elementary-unit stage of the RECS sample design, field workers compile lists of all housing units in the listing segments selected during the secondary stage. From these lists, a penultimate cluster of housing units is selected for each listing segment, and the ultimate cluster is selected from the penultimate cluster.

When an SSU is selected for the first time, the elementary-unit stage must be developed from the beginning. This includes the field-listing procedures, the penultimate cluster selection procedures, and the interviewer assignment procedures. For RECS cycles between major revisions, the elementary-unit stage may or may not be redeveloped, and in fact, with the majority of SSU's, it is not. Instead, the listing of housing units in the penultimate cluster is updated to account for new construction, demolitions, and other changes in the housing stock. If properly implemented, these new construction update procedures should not produce biased results, but rather, may yield a high variance for new home statistics. In addition, the provision of a longitudinal panel ensures that the ultimate clusters from one-half of the SSU's will be carried over from one RECS to the next. In the 1980 through 1990 RECS, these new construction update procedures were coordinated with the design of the longitudinal panel in order to keep the sample current, as well as to maintain continuity between RECS cycles.

The 1993 RECS included a new construction supplement, the effect of which, was to sample recently built housing units at a higher rate than older housing units. The sample design for this supplement involved both the secondary stage and the elementary-unit stage.

The elementary-unit stage is also designed to accommodate a low-income oversample, as well as to handle the existence of additional housing units discovered during the interviews.

Basic Elementary-Unit Stage

The following decisions are made at the elementary-unit stage:

- The instructions to the field workers for the field-listing procedures
- The maximum size for penultimate clusters
- How to select the penultimate clusters from the housing units found during the field-listing procedures
- The number of housing units to select from each penultimate cluster to form the ultimate clusters
- How to sample the housing units from the penultimate clusters for the ultimate clusters
- How to handle housing units in the penultimate clusters that were not listed during the field-listing procedures.

Instructions for Field Workers (Definition of Housing Unit)

The field workers are provided a map of the listing segment and are instructed to list all housing units they discover. These instructions define a housing unit[15] as follows:

> A house, an apartment, a group of rooms, or a single room if it is either occupied, or intended for occupancy, as separate living quarters by a family, an individual, or a group up to nine unrelated persons. Separate living quarters means the occupant(s): (1) live and eat separately from other persons in the house or apartment and (2) have direct access from the outside of the building or through a common hall -- that is, entry or exit can be made without entering someone else's living quarters. Housing units do not include group quarters such as prisons or nursing homes where ten or more unrelated persons live. A common dining area used by residents is an indication of group quarters. Hotel and motel rooms are considered housing units if occupied as the usual or permanent place of residence.

Appendix B contains the detailed instructions given to field workers to help them determine what constitutes a housing unit.

Maximum Size of Penultimate Clusters

The SSU's are typically larger in size than is desired for the penultimate clusters. The larger size of the SSU's makes the new construction update procedures used between major design years more effective, but using the full SSU adds costs in other areas. However, since the penultimate clusters are subsets of the SSU's, their use, for example, in computerizing and later rechecking the housing lists, reduce costs.

In the 1984 design, the maximum size for a penultimate cluster was 25 housing units. For the 1993 design this was increased to 50 housing units for the core SSU's, and 96 units for the new construction supplement of SSU's.

Selection of Penultimate Clusters

The penultimate clusters represent a compact geographic area within the SSU's, which are defined by the addresses of the housing units they contain. This means that a set of 25, 50, or 96 consecutive addresses define a compact geographic area or penultimate cluster. The penultimate clusters are established by selecting a single housing unit from the complete list of SSU housing units, and then defining this cluster as the next 25, 50, or 96 housing units which follow the single unit selected. The housing unit list is created in a circular manner. This refers to the process of returning to the beginning of the list, when its end is reached, in order to obtain the remaining units for the penultimate cluster.

Number of Observations in an Ultimate Cluster

The approximate number of housing units per ultimate cluster has consistently been close to four or five for each RECS. This corresponds to the results of the optimization analysis, which has called for a smaller number of observations in the ultimate clusters. The divergence between the actual number used and the

[15]This definition is taken from the *1993 RECS Instructions for Interviewers'* Manual--Glossary, p. 8.

number specified by the optimization analysis is discussed in Chapter 4. The average number of observations in the ultimate clusters for each cycle of RECS, is increased or decreased to reflect the target number of completed observations.

Selection of Housing Units for the Ultimate Clusters

The housing units in the ultimate clusters are selected from those in the penultimate clusters using simple random sampling without replacement. In the 1980 through 1990 RECS, the housing units in the penultimate clusters were not divided into strata prior to the selection of the ultimate clusters. Consequently, within SSU's, the selection probabilities were not dependent on housing-unit characteristics, but did vary between SSU's according to the number of available housing units in the penultimate cluster, and the target number of housing units in the ultimate cluster. This target number was higher when the SSU's were designated as part of a low-income oversample. In the 1993 RECS, the housing units in the penultimate clusters were divided into two strata using the estimated age of the housing units prior to the selection of the ultimate clusters. The housing units in the stratum composed of units judged by the field workers to be newly constructed were sampled at a higher rate than those in the stratum composed of older units.

The number of housing units in the penultimate cluster that are available for selection to the ultimate cluster, varies according to the status of the corresponding SSU and listing segment. If the listing segment is to be used for the first time, all housing units in the penultimate cluster are available. If the SSU is part of the longitudinal panel (see discussion below), the available housing units are from an ultimate cluster chosen in a previous RECS. (An ultimate cluster used in a longitudinal panel is generally the equivalent of one used in a previous RECS.) If a SSU (or listing segment) is designated as part of the incoming panel, only those housing units in the penultimate cluster that have not been selected for a previous RECS are available.

In previous RECS cycles, there have been listing segments where the target size of an ultimate cluster exceeds the number of available housing units. This has occurred in two situations. In the first situation, when a SSU is part of the incoming panel (see below), and its corresponding penultimate cluster has been used in several previous RECS samples, it is possible that previous RECS cycles have used up almost all of the housing units in the penultimate cluster. In this case, those housing units that have not been used earlier are placed in the ultimate cluster, and the remaining housing units needed to complete the ultimate cluster are selected from those used in previous surveys.

A second situation occurs when a SSU is part of a longitudinal panel, and the target number of housing units for the ultimate cluster exceeds the number used earlier when the SSU was part of the incoming panel. In this case, the additional housing units needed for the ultimate cluster are selected from those in the penultimate cluster, which have not been used in a previous RECS.

Treatment of Housing Units Discovered During the Interviews

Even with the use of field workers, the list of housing units in the listing segments is incomplete. Some housing units are missed. Many missed units are separate living quarters within structures that contain other housing units (e.g., a basement apartment rented to an unrelated individual), or houses that are hidden from view.

Housing units that are missed during the listing phase are sometimes discovered during the RECS interviews. If a housing unit is discovered during these interviews, there are rules to follow for determining if it should be included in the ultimate cluster. For example, if a housing unit that is selected for the ultimate cluster happens to be two units, then both units are included. Also, if an unlisted single-family home is found and its listing would immediately follow a housing unit in the ultimate cluster, it too is added to the ultimate cluster. On the other hand, if the unlisted home follows a housing unit that is not in the ultimate cluster, then

this home is not added to the ultimate cluster. Additional rules handle more complicated situations, such as cases where a set of four or more units have been left off the housing-unit list.

Longitudinal and Incoming Panels

The provision for a longitudinal panel was incorporated into the sample design for the 1980 RECS. In particular, the 1980 RECS sample design divided the SSU's selected during the secondary stage into four panels. Each of the four panels could be used as a national sample by itself. This feature was incorporated into the RECS sample design for the following reasons:

- The RECS could be fielded every 6 months, by interviewing a sample of housing units from two of the panels

- With the 6-month rotation, each housing unit would be interviewed twice. The original plans called for each 6-month cycle of RECS to consist of two panels: one panel where the housing units are interviewed for the first time, and another panel in which the housing units were interviewed 2 years earlier.

The RECS was never fielded every 6 months. Therefore, the 1980 RECS included all four panels and all of the housing units were interviewed for the first time. This was repeated in the 1981 RECS. Beginning with the 1982 RECS, the panels were divided into incoming panels and longitudinal panels. The housing units in the incoming panel were selected for the first time, while the housing units in the longitudinal panel had either been selected for the 1980 RECS, or added to the sample as part of the new construction update. Hence, slightly less than one-half of the housing units for the 1982 RECS had also been selected for the 1980 RECS. This longitudinal rotation of the sample was made a part of the 1984, 1987, and the 1990 RECS.

In the 1984 RECS, it was necessary to use the longitudinal panel from the 1980 design. The incoming panel, however, was developed in the 1984 design. This complication added higher costs to the 1984 RECS because it meant working in both new and old PSU's.

The 1993 RECS did not include a longitudinal panel. One reason for its elimination was to avoid the extra costs like those incurred in the 1984 RECS.

In the 1982 RECS, the longitudinal panel included housing units that were selected for the 1980 RECS. In the 1984, 1987 and 1990 RECS, the longitudinal panel included housing units that were initially chosen in the previous RECS (i.e., in the 1982, 1984, and 1987 RECS, respectively).

The primary objective of the longitudinal rotation plan (or longitudinal sample design), was to provide a subsample where changes occurring in the same group of housing units during the period between two RECS data-collection cycles could be analyzed. The period was 2 years for the 1982 and 1984 RECS and 3 years for the 1987 and 1990 RECS. This objective was accomplished using rotation groups. Systematic random procedures were used to divide the total set of SSU's into four subsamples, designated in Table 11 as C, D, E, and F.

Table 11. Overview of Longitudinal Sample Design for the 1982, 1984, 1987, and 1990 RECS

Rotation Group	1982	1984	1987	1990
C	R	S[a]	R	N
D	R	N[a]	R	S
E	S	R[b]	N[a]	R
F	N	R[b]	S[a]	R

[a] PSU's selected during the primary stage for the 1984 redesign were used for the first time.
[b] PSU's selected during the primary stage for the initial 1980 design were used for the returning panels for the 1984 RECS.
R = Housing units return from preceding survey.
S = Selected housing units from the same penultimate clusters that had been used in the preceding survey.
N = Selected new listing segments from the SSU's used in the preceding survey.
Source: Energy Information Administration, Office of Energy Markets and End use, the 1982, 1984, 1987, and 1990 Residential Energy Consumption Surveys.

In the 1990 RECS, for example, Rotation Groups E and F were designated as those returning groups, originally sampled in the previous RECS, which were to be interviewed again. Groups C and D on the other hand, made up the incoming or new groups, which were included for the first time in the 1990 RECS sample. Procedures for updating the sample for new construction and for other changes in the housing unit stock, were incorporated so that each rotation group, as well as the total RECS sample, represented a probability sample of the survey population.

The sample design for the 1993 RECS did not include a longitudinal panel, though one may be incorporated into the sample design for the 1996 and/or 1999 RECS. It would be possible to divide each of the ultimate clusters used for the 1993 RECS into two panels, and institute the longitudinal design at the elementary-unit level. Alternatively, the SSU's could be divided into two panels. (Each PSU for the 1993 design contains an even number of SSU's.) The longitudinal design could then be instituted at the SSU level, as was done for the 1980 through 1990 RECS. The exact form of a longitudinal sample design for the 1996 and/or 1999 RECS will be decided in the future. One complicating factor in developing a longitudinal panel for the 1996 RECS, is that the 1993 RECS obtained its SSU's from the second stages of both the 1984 and 1993 designs. Hence, some PSU's in the 1996 RECS would use the returning panel from 1984 SSU's, and the incoming panel from 1993 SSU's. The need to work with two sets of SSU's in some PSU's will add higher costs to any proposal to incorporate a longitudinal panel into the 1996 RECS sample design.

New Construction Update Procedures

The new construction procedures for the 1982 through 1990 RECS vary by rotation group. The procedures for the returning rotation groups were different than those for the incoming rotation groups. In addition, the procedures differed between the two incoming rotation groups. The new construction procedures for the 1990 RECS will be used to describe the procedures in general.

New Construction Updates for Returning Rotation Groups

Groups E and F made up the returning rotation groups for the 1990 RECS. The general plan for the SSU's in the returning groups, was to attempt to conduct interviews in the housing units taken from the ultimate clusters in the 1987 RECS, as well as in a sample of newly constructed units. It should be noted that these ultimate clusters in the 1987 RECS also included housing units that had been vacant when the interviews were first attempted, as well as housing units where interviews were attempted unsuccessfully (refusals, not-at-home, etc.).

In mid-1990, before the 1990 RECS interview phase was started in Rotation Groups E and F, field workers made visits to their respective penultimate clusters. During these visits, the 1987 housing units listings were checked and updated for missed units, new construction, demolition, and structure conversion (i.e., conversion from one use to another).

A sample of newly constructed units, units converted from nonresidential uses to residential units, and units missed during the 1987 RECS field-listing procedures, were added to the ultimate clusters for Rotation Groups E and F. These units were sampled from the updated field listings at the 1990 RECS sampling rate.

New Construction Updates for New Rotation Groups

Groups C and D formed the new rotation groups for the 1990 RECS. Prior to selecting the ultimate clusters for the 758 SSU's in these groups, a new construction update procedure was utilized. The update procedure started with a canvass (primarily by telephone) of local information sources, such as building-permit-issuing agencies, zoning boards, and tax offices. The objective was to determine whether significant new construction--defined as groups of 25 or more new housing units--had occurred within the SSU since 1984. In this canvass, significant new construction was found in 197 of the 758 SSU's. The rough-counting procedures, segment formulation procedures, and segment selection procedures of the secondary stage, as well as the entire primary stage, were repeated for these 197 SSU's.

In the SSU's where no significant new construction was found, procedures differed in Rotations Groups C and D. In mid-1990, field workers also made visits to the penultimate clusters in the SSU's in Rotation Group D. Like the listings checks carried out for Rotation Groups E and F, in these visits, workers checked and updated the 1987 housing units listings for missed units, new construction, demolitions, and structure conversions. In the SSU's in Rotation Group D, housing units for the 1990 RECS sample were selected from among those *not* selected in the earlier RECS.

For SSU's in Rotation Group C, a new listing segment was selected for the 1990 RECS using the existing RECS rough counts. The elementary-unit stage was redeveloped for these new listing segments.

New Construction Update for 1993 RECS

Initial plans for the 1993 design included the complete revision of the secondary stage and, subsequently, the development of the elementary-unit stage for those SSU's selected in the secondary stage. While initially, the new construction update procedures were not needed, the budget restrictions required some PSU's to take their secondary stages from the 1984 design. For these PSU's, the new construction update procedures were used and were the same as those used in prior RECS for the new rotation groups.

New Construction Update for 1996 RECS

New construction update procedures are included in the plans for the 1996 RECS. The exact form of these procedures has not been completely specified. It is anticipated that the updates will cover not only the sampled SSU's but, also additional Census Blocks that adjoin the sampled SSU's.

New Construction Supplement for the 1993 RECS

An additional feature of the 1993 RECS design, is an oversample of newly built housing units. The design of the secondary stage for the 1993 RECS calls for 1,460 SSU's in the core sample, and 150 supplemental SSU's from those Census tracts or block groups with a high percentage of new housing units (i.e., constructed

in the 6-year period prior to the 1990 Census). In addition, housing units judged to be new by field workers during the field-listing procedures will be sampled at a higher rate than those judged to be older. The oversample was included as part of the 1993 RECS sample design because of the importance of obtaining accurate data on the energy characteristics of new housing units.

One of the reasons for the major revisions in the sample design prior to the 1984 and 1993 RECS, was to update the sample design for population changes measured by the 1980 and 1990 Censuses. Areas with a large population increase should have a much higher incidence of new residential construction, than areas with a smaller population increase, or a decrease in population.

Low-Income Supplement

An additional feature of the 1981, 1984, 1987, and 1993 RECS sample designs, is an oversample of housing units in SSU's determined by the interviewers to be low-income areas. Because of the low number of low-income housing units in the core sample whose main space-heating fuel is electricity, fuel oil, liquefied petroleum gas, or kerosene, housing units in low-income SSU's where the main space-heating fuel was something other than natural gas were sampled at even an higher rate. This low-income supplement was funded by the Administration for Children and Families of the U.S. Department of Health and Human Services (HHS). Data collected during the RECS on low-income households are used by HHS to administer the Low-Income Home Energy Assistance Program.

Summary of Features for the Elementary-Unit Stage

Table 12 lists the sample size, the presence of a longitudinal component, the presence of a low-income supplement, and the presence of a new construction supplement.

Table 13 summarizes the new construction update/oversample procedures for the various cycles of RECS.

Table 12. Features of RECS Surveys

Survey	Number of Observations	Longitudinal Component	Low-Income Supplement	New Construction Supplement
1978 NIECS	3,842	None	None	None
Screener (1979)	3,064	None	None	None
1980 RECS	5,804	None	None	None
1981 RECS	6,269	None	Approximately 18.4 percent of observations	None
1982 RECS	4,724	2 of 4 panels carried over from 1980 RECS	None	None
1984 RECS	5,682	2 of 4 panels carried over from 1982 RECS	Approximately 19.3 percent of observations	None
1987 RECS	6,229	2 of 4 panels carried over from 1984 RECS	Approximately 17.5 percent of observations	None
1990 RECS	5,095	2 of 4 panels carried over from 1987 RECS	None	None
1993 RECS	Approximately 7,050	None	Approximately 850 observations out of 7,050 observations	Approximately 1,200 observations out of 7,050 observations

Source: Energy Information Administration, 1978, 1979, 1980, 1981, 1982, 1984, 1987, 1990, and 1993 Residential Energy Consumption Surveys.

Table 13. New Construction Updates

Survey	Half-Open Interval Procedures Used	Local Officials Contacted and SSU's Relisted If Large Amount of New Construction in SSU	Other New Construction Update Procedures
NIECS (1978) and Screener (1979)	Yes	No	Selected 56 additional SSU's
1980 RECS	Yes	No	
1981 RECS	Yes	Yes	
1982 RECS	Yes	Yes	
1984 RECS	Yes	Yes, for SSU's retained from 1980 design	
1987 RECS	Yes	Yes	
1990 RECS	Yes	Yes	
1993 RECS	Yes	Yes, for SSU's retained from 1984 design	

Source: Energy Information Administration, 1978, 1979, 1980, 1981, 1982, 1984, 1987, 1990, and 1993 Residential Energy Consumption Surveys.

8. Reasons for Revising RECS Sample Design

The RECS sample design will need to be updated/revised/redesigned periodically. The next major revision is scheduled to take place when the data for the 2000 Census become available. If major additional demands are placed on the RECS design, such as the capacity to produce State-level estimates, then it may need to be substantially revised in order to meet these demands prior to the post-2000 Census redesign. Even if the primary stage is not revised, the secondary and elementary-unit stages will need to be periodically revised.

Data from the Decennial Censuses are extremely useful in the development and revision of the sample design for household surveys. Traditionally, the sample design for household surveys is updated every 10 years when the Census data becomes available. The Census data are used to develop improved definitions of PSU's and improved groupings of the PSU's into strata (both in greater MOS equality and in greater homogeneity of energy characteristics within strata). Similar measures also are possible in the secondary stage by improving the definition and stratification of SSU's. The use of new Census data to determine the MOS for PSU's and SSU's will reduce the variance of estimates derived from housing-unit surveys, particularly for estimates concerning new construction.

There are two major reasons for a regular decennial redesign of the RECS sample design:

- Outdated survey materials and definitions
- Uneven population growth.

Outdated Survey Materials and Definitions

The following survey materials and definitions may become outdated:

- Definitions of PSU's, SSU's, and listing segments
- Maps of SSU's and listing segments
- Lists of housing units in listing segments.

The definitions of metropolitan PSU's make use of MSA boundaries. These definitions are periodically revised by the Office of Management and Budget. The boundaries of MSA PSU's are brought up to date during the decennial revisions of the RECS sample design.

The SSU definitions make use of Census blocks and block groups. The listing segment for an SSU is either equal to the full SSU or only part of the SSU. Maps of the listing segment are used by the field workers to define the area in which housing units are to be listed, and are also used by interviewers to locate the housing units in the ultimate cluster. The construction of new roads, new landmarks, and new housing units entails that these maps will become outdated over time.

The minimum size of the listing segments was chosen, in part, to ensure that the segment will contain enough housing units so that each unit will be selected for no more than two cycles of the RECS. If new listing segments are not periodically chosen, the housing-unit lists will be exhausted.

Some listing segments will be contained in a single apartment complex. Housing units in this complex will be sampled for each RECS as long as the listing segment is used.

Uneven Population Growth

While the national population grew by 9.8 percent from 1980 to 1990, there has been much variation in the rates of population change within and between Census divisions, States, and counties. This percentage of population change between geographic locations, can be illustrated by using the 1980 and 1990 Censuses.

Uneven population growth implies that both the primary stage and secondary stage of the design need to be periodically revised or updated. The extent and effect of uneven population growth will be briefly discussed below.

Population Trends by Census Division and State

The U.S. population is regularly enumerated in a Decennial Census. Table 14 shows the national population and the population by State and by Census division for the 1990 and 1980 Censuses. It also indicates the percentage change during these years.

All nine of the Census divisions showed a growth in population. This population increase ranges from 0.8 percent for the East North Central Census Division, to 23.0 percent for the Pacific Census Division. The two fastest growing divisions between 1980 and 1990 were the Pacific and the Mountain Census Divisions, both of which are in the West Census Region. The slowest growing divisions include the Middle Atlantic, the East North Central, and the West North Central Census Divisions.

Within Census divisions, the percentage of population increase varies by State. To illustrate an extreme, within the Mountain Census Division, Wyoming's population decreased by 3.4 percent, while Nevada's population increased by 50.1 percent in the same 1980 to 1990 period. Similarly, within the South Atlantic Census Division, West Virginia's population decreased by 8.0 percent, while Florida's population increased by 32.7 percent. The range in population growth was smallest in the East North Central Census Division. In this division, Illinois' population increased by less than 0.05 percent, while Wisconsin's population increased by 4.0 percent.

Nevada had the largest percentage of population increase (50.1 percent) of any State in the Union. While the populations of West Virginia and the District of Columbia decreased by the largest percentage: 8.0 percent and 4.9 percent, respectively.

Population Trends by County

Within States, some counties lost, while other counties gained in population. Table 15 indicates by State and by Census division, the number of counties or equivalent areas, which lost or gained in population during the 1980's. Table 16 shows the minimum and maximum percentage change within counties, State, Census division and, national total.

Of the 3,141 counties or equivalent areas nationwide, 1,431 or 46 percent of these experienced a decrease in population from 1980 to 1990. The populations of Platte County in Wyoming, and Lake County in Colorado decreased by the largest percentage (32 percent). Flagler County in Florida showed the largest percentage of population increase (163 percent).

Table 14. Population and Percent Increase in Population by State, Census Division, and National Total

Census Regions and Divisions	Population (thousands)		Percent Increase in Population
	1990	1980	
National Total	248,710	226,546	9.8
New England	13,207	12,348	7.0
Maine	1,228	1,125	9.2
New Hampshire	1,109	921	20.5
Vermont	563	511	10.0
Massachusetts	6,016	5,737	4.9
Rhode Island	1,003	947	5.9
Connecticut	3,287	3,108	5.8
Middle Atlantic	37,602	36,787	2.2
New York	17,990	17,558	2.5
New Jersey	7,730	7,365	5.0
Pennsylvania	11,882	11,864	0.1
East North Central	42,009	41,682	0.8
Ohio	10,847	10,798	0.5
Indiana	5,544	5,490	1.0
Illinois	11,431	11,427	0.0
Michigan	9,295	9,262	0.4
Wisconsin	4,892	4,706	4.0
West North Central	17,660	17,183	2.8
Minnesota	4,375	4,076	7.3
Iowa	2,777	2,914	-4.7
Missouri	5,117	4,917	4.1
North Dakota	639	653	-2.1
South Dakota	696	691	0.8
Nebraska	1,578	1,570	0.5
Kansas	2,478	2,364	4.8
South Atlantic	43,567	36,959	17.9
Delaware	666	594	12.1
Maryland	4,781	4,217	13.4
District of Columbia	607	638	-4.9
Virginia	6,187	5,347	15.7
West Virginia	1,793	1,950	-8.0
North Carolina	6,629	5,882	12.7
South Carolina	3,487	3,122	11.7
Georgia	6,478	5,463	18.6
Florida	12,938	9,746	32.7
East South Central	15,176	14,666	3.5
Kentucky	3,685	3,661	0.7
Tennessee	4,877	4,591	6.2
Alabama	4,041	3,894	3.8
Mississippi	2,573	2,521	2.1
West South Central	26,703	23,747	12.4
Arkansas	2,351	2,286	2.8
Louisiana	4,220	4,206	0.3
Oklahoma	3,146	3,025	4.0
Texas	16,987	14,229	19.4

Table 14. Population and Percent Increase in Population by State, Census Division, and National Total (Continued)

Census Regions and Divisions	Population (thousands)		Percent Increase in Population
	1990	1980	
Mountain	13,659	11,373	20.1
Montana	799	787	1.6
Idaho	1,007	944	6.7
Wyoming	454	470	-3.4
Colorado	3,294	2,890	14.0
New Mexico	1,515	1,303	16.3
Arizona	3,665	2,718	34.8
Utah	1,723	1,461	17.9
Nevada	1,202	800	50.1
Pacific	39,127	31,800	23.0
Washington	4,867	4,132	17.8
Oregon	2,842	2,633	7.9
California	29,760	23,668	25.7
Alaska	550	402	36.9
Hawaii	1,108	965	14.9

Source: Statistical Abstract of the United States 1993, U.S. Department of Commerce, Bureau of the Census. (Table No. 31, Page 28.)

Table 15. Number of Counties Losing or Gaining Population by State, Census Division, and National Total (Continued)

Census Regions and Divisions	Number of Counties						
		Losing Population		Gaining Population or No Change			
	Total	More than 10 % Loss	0 % Loss to 10 % Loss	No Change to 10 % Gain	11 % Gain to 20 % Gain	21 % Gain to 40 % Gain	More Than 40 % Gain
National Total	3,141	457	974	952	376	262	120
New England	67	0	3	33	24	7	0
Maine	16	0	1	7	8	0	0
New Hampshire	10	1	1	3	5	0	0
Vermont	14	0	0	7	7	0	0
Massachusetts	14	0	1	10	1	2	0
Rhode Island	5	0	0	4	1	0	0
Connecticut	8	0	0	4	4	0	0
Middle Atlantic	150	5	52	73	15	4	1
New York	62	0	16	43	2	1	0
New Jersey	21	0	4	9	6	2	0
Pennsylvania	67	5	32	21	7	1	1
East North Central	437	25	193	170	42	7	0
Ohio	88	4	35	40	8	1	0
Indiana	92	0	50	36	5	1	0
Illinois	102	18	63	14	6	1	0
Michigan	83	3	26	35	15	4	0
Wisconsin	72	0	19	45	8	0	0
West North Central	618	238	215	110	36	15	4
Minnesota	87	23	26	26	6	4	2
Iowa	99	50	41	7	1	0	0
Missouri	115	21	31	36	17	8	2
North Dakota	53	34	13	5	1	0	0
South Dakota	66	30	22	9	5	0	0
Nebraska	93	43	40	8	2	0	0
Kansas	105	37	42	19	4	3	0
South Atlantic	591	40	140	165	107	84	55
Delaware	3	0	0	0	3	0	0
Maryland	24	0	3	5	6	8	2
District of Columbia	1	0	1	0	0	0	0
Virginia	136	9	46	34	21	17	9
West Virginia	55	21	23	6	4	1	0
North Carolina	100	0	20	42	28	7	3
South Carolina	46	5	7	23	6	3	2
Georgia	159	5	38	49	30	22	15
Florida	67	0	2	6	9	26	24
East South Central	364	121	147	156	22	16	2
Kentucky	120	9	57	45	8	1	0
Tennessee	95	1	26	54	6	7	1
Alabama	67	3	32	24	4	3	1
Mississippi	82	8	32	33	4	5	0
West South Central	470	69	148	140	55	42	16
Arkansas	75	10	25	29	8	3	0
Louisiana	64	6	31	21	3	3	0
Oklahoma	77	15	32	23	5	2	0
Texas	254	38	60	67	39	34	16

Table 15. Number of Counties Losing or Gaining Population by State, Census Division, and National Total (Continued)

Census Regions and Divisions	Number of Counties						
		Losing Population		Gaining Population or No Change			
	Total	More than 10 % Loss	0 % Loss to 10 % Loss	No Change to 10 % Gain	11 % Gain to 20 % Gain	21 % Gain to 40 % Gain	More Than 40 % Gain
Mountain	281	55	60	62	50	30	24
Montana	57	21	19	8	9	0	0
Idaho	44	7	11	17	6	3	0
Wyoming	23	9	5	6	1	1	1
Colorado	63	11	16	9	12	9	6
New Mexico	33	2	6	9	11	2	3
Arizona	15	1	0	1	5	4	4
Utah	29	4	3	10	5	5	2
Nevada	17	0	0	2	1	6	8
Pacific	163	4	16	43	25	57	18
Washington	39	1	5	18	5	10	0
Oregon	36	3	10	14	7	2	0
California	58	0	0	7	11	30	10
Alaska	25	0	0	3	2	13	7
Hawaii	5	0	1	1	0	2	1

Source: U.S. Bureau of the Census, *County Populations, 1990 and 1980 Censuses Ranked By 1990 Population Within States*, U.S. Government Printing Office (Washington, DC), 1991.

Table 16. Percent Increase in Population, Smallest Percent Increase in Population for a County and Largest Percent Increase in Population for a County by State, Census Division, and National

Census Regions and Divisions	Percent Increase in Population for State, Census Division, or Nation	Smallest Percent Increase in Population for a County	Largest Percent Increase in Population for a County
National Total	9.8	-32.0	163.0
New England	7.0	-4.8	30.2
Maine	9.2	-4.8	18.2
New Hampshire	20.5	-0.9	29.2
Vermont	10.0	1.5	17.7
Massachusetts	4.9	-4.0	30.2
Rhode Island	5.9	4.1	17.9
Connecticut	5.8	2.5	12.1
Middle Atlantic	2.2	-11.4	53.1
New York	2.5	-4.6	25.9
New Jersey	5.0	-8.6	25.2
Pennsylvania	0.1	-11.4	53.1
East North Central	0.8	-16.5	32.8
Ohio	0.5	-13.9	24.3
Indiana	1.0	-10.0	32.8
Illinois	0.0	-16.5	23.9
Michigan	0.4	-13.5	29.5
Wisconsin	4.0	-8.6	16.5
West North Central	2.8	-23.8	47.7
Minnesota	7.3	-20.1	41.7
Iowa	-4.7	-17.1	17.6
Missouri	4.1	-20.5	47.7
North Dakota	-2.1	-23.8	16.6
South Dakota	0.8	-20.2	15.6
Nebraska	0.5	-22.1	19.3
Kansas	4.8	-18.9	38.8
South Atlantic	17.9	-29.4	163.0
Delaware	12.1	11.0	15.5
Maryland	13.4	-7.0	58.0
District of Columbia	-4.9	-4.9	-4.9
Virginia	15.7	-22.9	80.3
West Virginia	-8.0	-29.4	26.7
North Carolina	12.7	-8.5	70.0
South Carolina	11.7	-7.2	42.0
Georgia	18.6	-22.1	125.6
Florida	32.7	-4.2	163.0
East South Central	3.5	-24.0	49.9
Kentucky	0.7	-12.7	25.6
Tennessee	6.2	-10.5	41.1
Alabama	3.8	-15.0	49.9
Mississippi	2.1	-24.0	29.7
West South Central	12.4	-30.2	91.1
Arkansas	2.8	-19.3	9.9
Louisiana	0.3	-20.5	30.3
Oklahoma	4.0	-19.6	31.8
Texas	19.4	-30.2	91.1

Table 16. Percent Increase in Population, Smallest Percent Increase in Population for a County and Largest Percent Increase in Population for a County by State, Census Division, and National (Continued)

Census Regions and Divisions	Percent Increase in Population for State, Census Division, or Nation	Smallest Percent Increase in Population for a County	Largest Percent Increase in Population for a County
Mountain	20.1	-32.0	140.1
Montana	1.6	124.7	17.7
Idaho	6.7	-27.5	37.7
Wyoming	-3.4	-32.0	43.7
Colorado	14.0	-32.0	140.1
New Mexico	16.3	-21.6	84.1
Arizona	34.8	-29.8	67.4
Utah	17.9	-19.7	86.3
Nevada	50.1	1.2	96.5
Pacific	23.0	-16.5	122.7
Washington	17.8	-13.2	37.9
Oregon	7.9	-16.5	26.7
California	25.7	1.5	58.5
Alaska	36.9	4.2	122.7
Hawaii	14.9	-9.7	41.7

[1]Negative number indicates a decrease in population.

Source: U.S. Bureau of the Census, *County Populations, 1990 and 1980 Censuses Ranked by 1990 Population Within State*, U.S. Government Printing Office (Washington, DC), 1991.

The West North Central Census Division had the highest percentage of counties that lost population (73 percent). The New England Census Division showed the lowest percentage (4 percent).

Among States, Iowa had the highest proportion of counties losing population (92 percent), followed by Nebraska (89 percent). Outside the Midwest, West Virginia showed 80 percent of its counties losing population. The District of Columbia, which can be treated as a single county-like area, experienced a 4.9 decrease in its population.

Conversely, all of the counties in Vermont, Rhode Island, Connecticut, Delaware, Nevada, Alaska, and California experienced an increase in population from 1980 to 1990. However, except for California, all of these States had a relatively small number of counties.

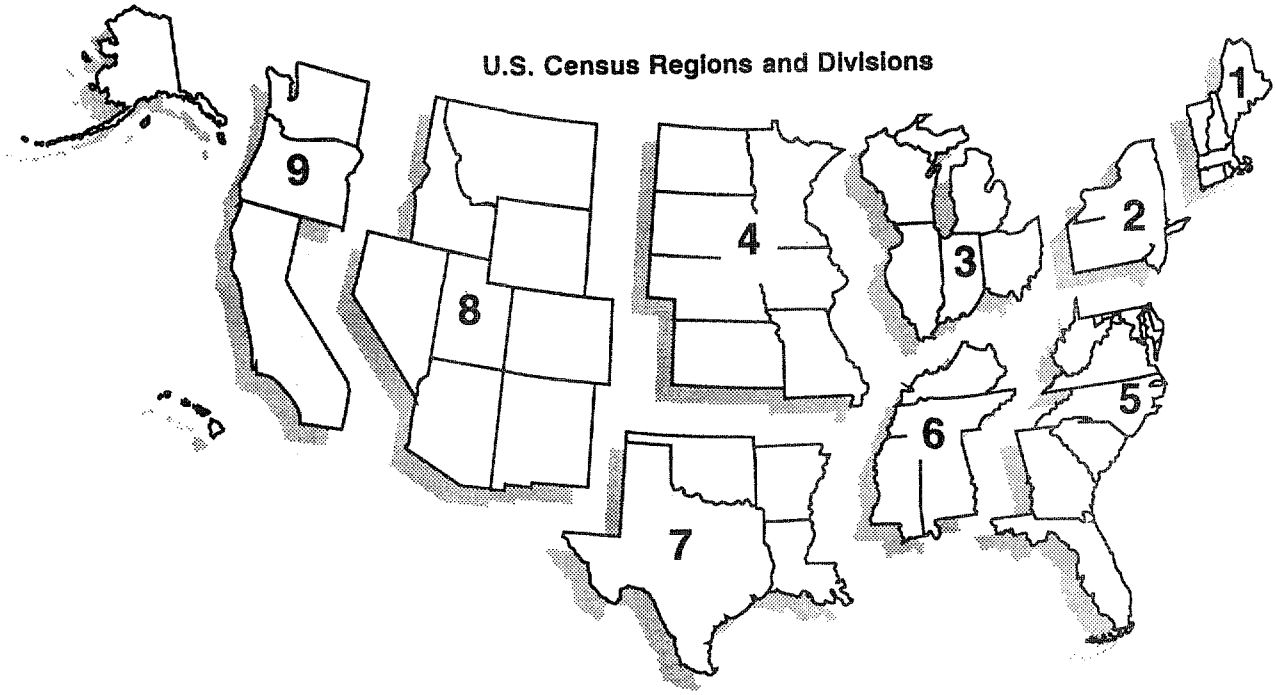
Tables 15 and 16 indicate that even within States, there is a wide range in the percentage of population change among counties. For example, of the 159 counties in Georgia, five of these experienced a 10 percent or more decrease, and 38 showed a 0 to 10 percent decrease. Yet, for 22 counties the population increased from 21 to 40 percent, and for 15 counties it increased by more than 40 percent.

When proceeding from larger to smaller segments of the population, the variability in percentage change between the segments increases. For example, the population change percentage between divisions ranged from a low of 0.8 percent increase (East North Central), to a high of 23.0 percent increase (Pacific). Between States the range was from an 8.0 percent decrease (West Virginia) to a 50.1 percent increase (Nevada), and between counties the range was from a 32 percent decrease to a 163 percent increase. At the Census tract level, a geographic location can experience a 100 percent decrease in population over a 10-year period if all of the housing units in it are demolished for commercial development (such as in the expansion of an airport or the demolition of a large apartment complex). Conversely, an agricultural area can experience a growth in population ranging from very few to hundred's, if over a 10-year period a major housing project is built.

The periodic redesign of the RECS is essential if accurate data on new housing units are to be obtained. As a design becomes older and, therefore, more out-of-date, its ability to support new construction estimates declines. The periodic redesign of all RECS stages is needed to account for the uneven nature of population growth. If the secondary stage and elementary-unit stage are not revised, only those new housing units in the existing sampled SSU's and listing segments will have a chance of being selected. The selection probabilities for PSU's and SSU's are based on either the population or the number of households at a set point in time. A nonuniform population increase will not produce biased results from keeping these PSU's and SSU's. However, it will imply that the variance of statistics will be larger than would be the case if the survey was redesigned. This will be particularly true for new housing units.

Appendix A

Census Divisions and Federal Regions Maps



Region 1
Northeast

Division 1
(New England)
Connecticut (CT)
Maine (ME)
Massachusetts (MA)
New Hampshire (NH)
Rhode Island (RI)
Vermont (VT)

Division 2
(Middle Atlantic)
New Jersey (NJ)
New York (NY)
Pennsylvania (PA)

Region 2
Midwest

Division 3
(East North Central)
Illinois (IL)
Indiana (IN)
Michigan (MI)
Ohio (OH)
Wisconsin (WI)

Division 4
(West North Central)
Iowa (IA)
Kansas (KS)
Minnesota (MN)
Missouri (MO)
Nebraska (NE)
North Dakota (ND)
South Dakota (SD)

Region 3
South

Division 5
(South Atlantic)
Delaware (DE)
District of Columbia (DC)
Florida (FL)
Georgia (GA)
Maryland (MD)
North Carolina (NC)
South Carolina (SC)
Virginia (VA)
West Virginia (WV)

Division 6
(East South Central)
Alabama (AL)
Kentucky (KY)
Mississippi (MS)
Tennessee (TN)

Division 7
(West South Central)
Arkansas (AR)
Louisiana (LA)
Oklahoma (OK)
Texas (TX)

Region 4
West

Division 8
(Mountain)
Arizona (AZ)
Colorado (CO)
Idaho (ID)
Montana (MT)
Nevada (NV)
New Mexico (NM)
Utah (UT)
Wyoming (WY)

Division 9
(Pacific)
Alaska (AK)
California (CA)
Hawaii (HI)
Oregon (OR)
Washington (WA)

Region 1
New England
Connecticut (CT)
Maine (ME)
Massachusetts (MA)
New Hampshire (NH)
Rhode Island (RI)
Vermont (VT)

Region 2
New York/New Jersey
New Jersey (NJ)
New York (NY)

Region 3
Mid Atlantic
Delaware (DE)
District of Columbia (DC)
Maryland (MD)
Pennsylvania (PA)
Virginia (VA)
West Virginia (WV)

Region 4
South Atlantic
Alabama (AL)
Florida (FL)
Georgia (GA)
Kentucky (KY)
Mississippi (MS)
North Carolina (NC)
South Carolina (SC)
Tennessee (TN)

Region 5
Midwest
Illinois (IL)
Indiana (IN)
Michigan (MI)
Minnesota (MN)
Ohio (OH)
Wisconsin (WI)

Region 6
Southwest
Arkansas (AR)
Louisiana (LA)
New Mexico (NM)
Oklahoma (OK)
Texas (TX)

Region 7
Central
Iowa (IA)
Kansas (KS)
Missouri (MO)
Nebraska (NE)

Region 8
North Central
Colorado (CO)
Montana (MT)
North Dakota (ND)
South Dakota (SD)
Utah (UT)
Wyoming (WY)

Region 9
West
Arizona (AZ)
California (CA)
Hawaii (HI)
Nevada (NV)

Region 10
Northwest
Alaska (AK)
Idaho (ID)
Oregon (OR)
Washington (WA)

Appendix B

Detailed Instructions to Field Workers on Housing Unit Definition[16]

A *housing unit* is a house, an apartment, a group of rooms, or a single room if it is occupied, or intended for occupancy, as separate living quarters by a family, an individual, or a group of up to nine unrelated persons. In general:

- a "one-family house" is one housing unit.
- a "two-family house" is two housing units
- each apartment in an apartment building is a housing unit
- a vacant house or apartment is a housing unit (because someone could live there in the future), and
- a store or business is not a housing unit, but ... an apartment over or behind a store or business is a housing unit.

Rules about *separate living quarters* may be needed to determine the number of housing units to be listed. Separate living quarters are those where:

- the occupant(s) live and eat separately from other persons in the building; and
- the occupant(s) have direct access from the outside of the building or through a common hall (that is, entry and exit can be made without entering someone else's living quarters).

Kitchen facilities are *not* required in order for separate living quarters to be defined as a housing unit.

Group Quarters - Living arrangements for institutional residents or inmates or for other groups of ten or more unrelated persons. Group quarters are *not* housing units, and you are not to conduct interviews in them. There are two general types:

- Group quarters are most frequently found in institutions. Examples are hospitals, nursing homes, military barracks, college dormitories, fraternity and sorority houses, convents, monasteries, and penal institutions.
- A house or apartment is considered to be group quarters if it is shared by the person in charge and nine or more *nonrelatives* (a total of ten or more persons). If there is no head of household or person in charge, the household or apartment is considered group quarters if it is shared by ten or more unrelated persons.

However, rooms or apartments within a institution that serve as the *permanent residence* of staff members *are* housing units if they satisfy the requirements of separate living quarters.

[16] This section is taken from the *1987 RECS Sample Design Procedures Manual*, pages F-10 to F-14.

Be alert for places that appear to be group quarters but do not contain enough people. For example, suppose that a convent turns out to have five nuns living in it. This convent would not be considered group quarters; it would be a housing unit.

Rooming and Boarding Houses - The rules apply here as applied to housing units/group quarters. The key items of information needed to determine whether these are housing units are (1) the number of residents and (2) whether some rooms, or groups of rooms, meet the requirements for separate living quarters.

Example 1: *Fewer than nine residents* unrelated to the person in charge, and separate rooms (or groups of rooms) do *not* meet conditions for separate living quarters.

This house would be treated as a single-housing unit.

Example 2: *Nine or more residents* unrelated to the person in charge (or a total of ten or more unrelated persons), and separate rooms (or groups of rooms) do *not* meet conditions listed for separate living quarters.

This house is considered a group quarters. No interview is to be conducted.

Example 3: Some rooms (or groups of rooms) meet the conditions for separate living quarters.

Consider each room (or group of rooms) that is *separate living quarters* as a separate housing unit.

Housing Units in Special Cases

Below are some examples of special situations that should be listed as housing units:

Basement or attic apartments which may occur in any type of structure.

Vacant houses or apartments (because they could be occupied in the future.)

Houses and apartments that are under construction - if they are likely to be completed and ready for occupancy by the time the interviews are started.

Hotel or motel rooms which are (1) occupied by permanent guests, or (2) occupied by employees who have no permanent residence elsewhere.

Rooms within group quarters or institutions (such as a fraternity house or dormitory) which serve as the *permanent residence* of a staff member or person in charge, and which satisfy the requirements of the housing-unit definition.

A room in a nonresidential structure -- if there is one room in a warehouse which the caretaker uses for his living quarters, such a room qualifies as a housing unit.

A mobile home trailer that is used as the permanent residence of occupants and not just as their vacation residence.

Work camps occupied by seasonal workers.

Seasonal dwellings, such as summer homes, resort cottages, or other part-time homes, are considered housing units.

Boats, tents, etc. Any such quarters that are occupied as someone's permanent residence are housing units. If they are used for vacation only, they are not housing units.

Not Classified as Housing Units in Special Cases

Below are some examples of situations when housing units should *not* be listed as such:

Group quarters--However, you cannot completely ignore such buildings, since some of them may contain the dwelling unit of a manager, janitor, etc. You should determine whether there are "hidden" housing units in the building.

Unoccupied buildings that have been *condemned* or that are being *demolished.*

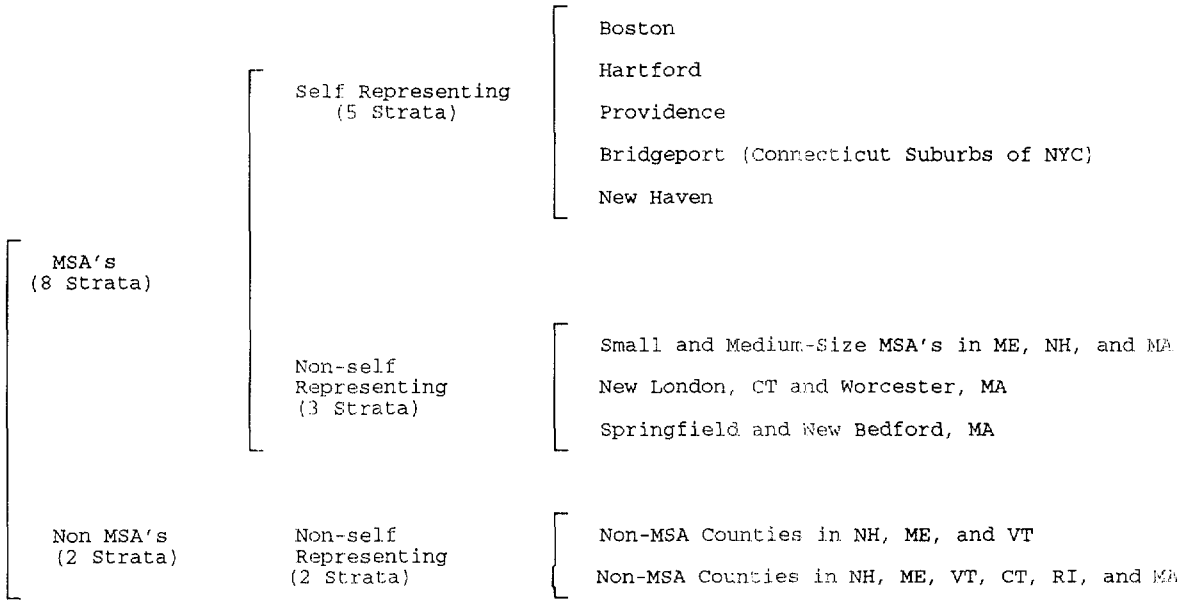
Places of business, such as stores, factories, etc.--but be sure to look for hard-to-find living quarters behind or above or inside business places.

Appendix C

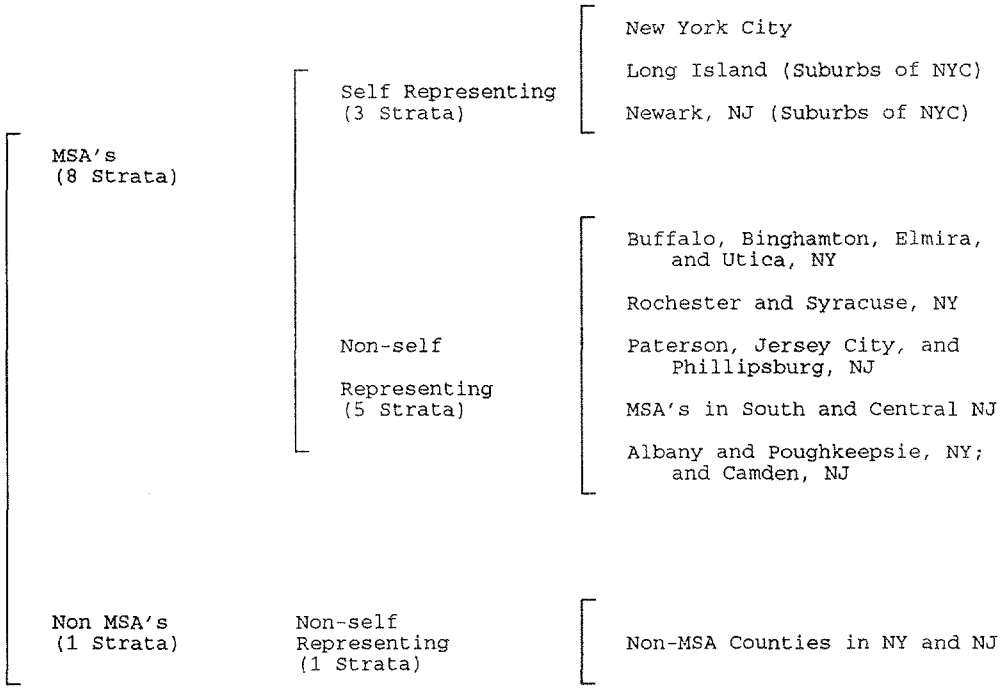
Strata Listings by Census Division

RECS 1980

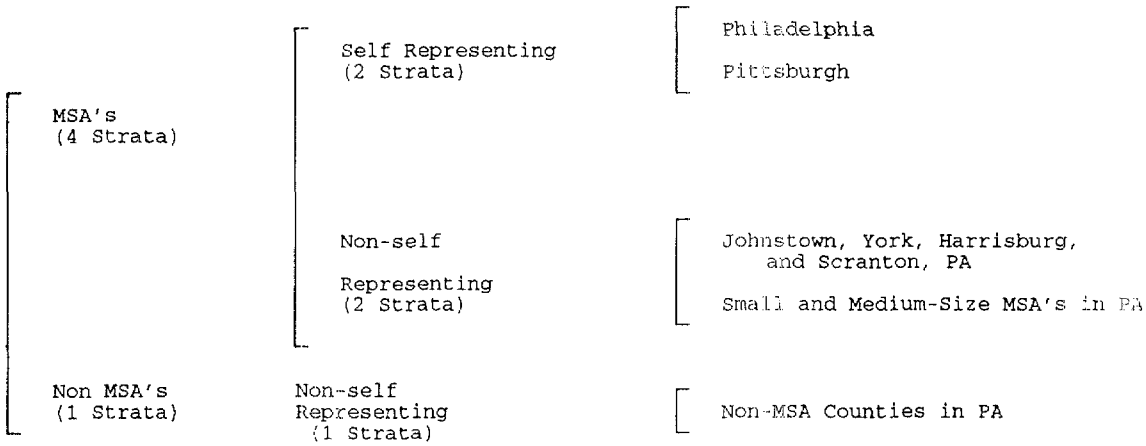
Intersection 1: Connecticut, Maine, Massachusetts, Rhode Island, and Vermont (10 Strata)



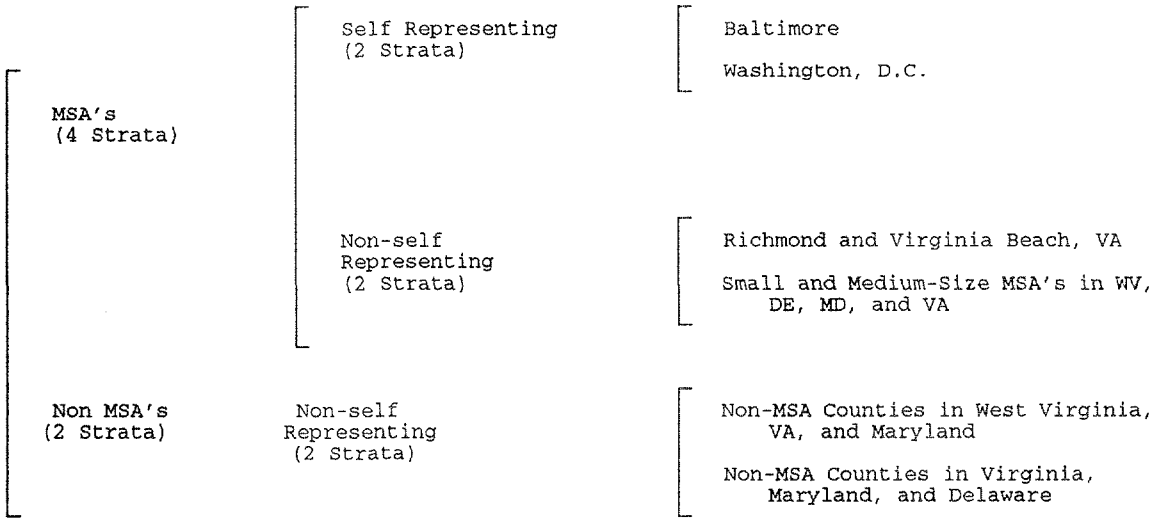
RECS 1980
Intersection 2: New York and New Jersey
(9 Strata)



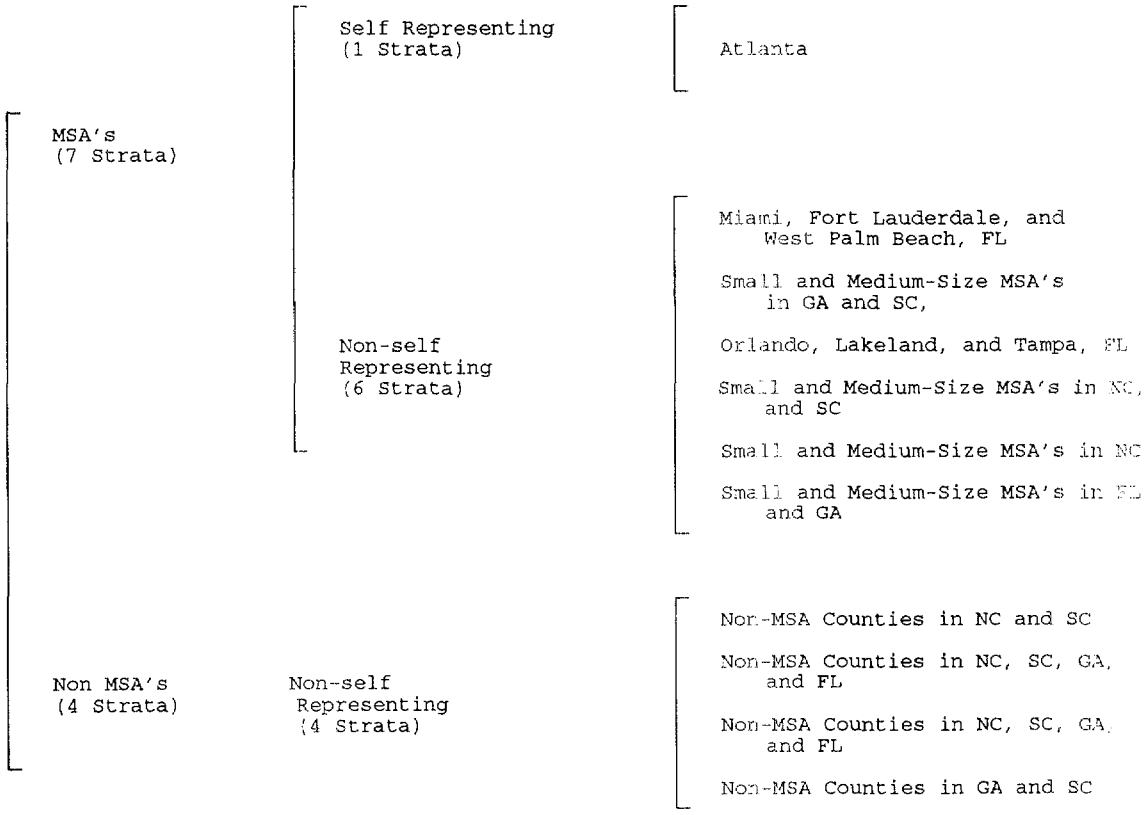
RECS 1980
Intersection 3: Pennsylvania
(5 Strata)



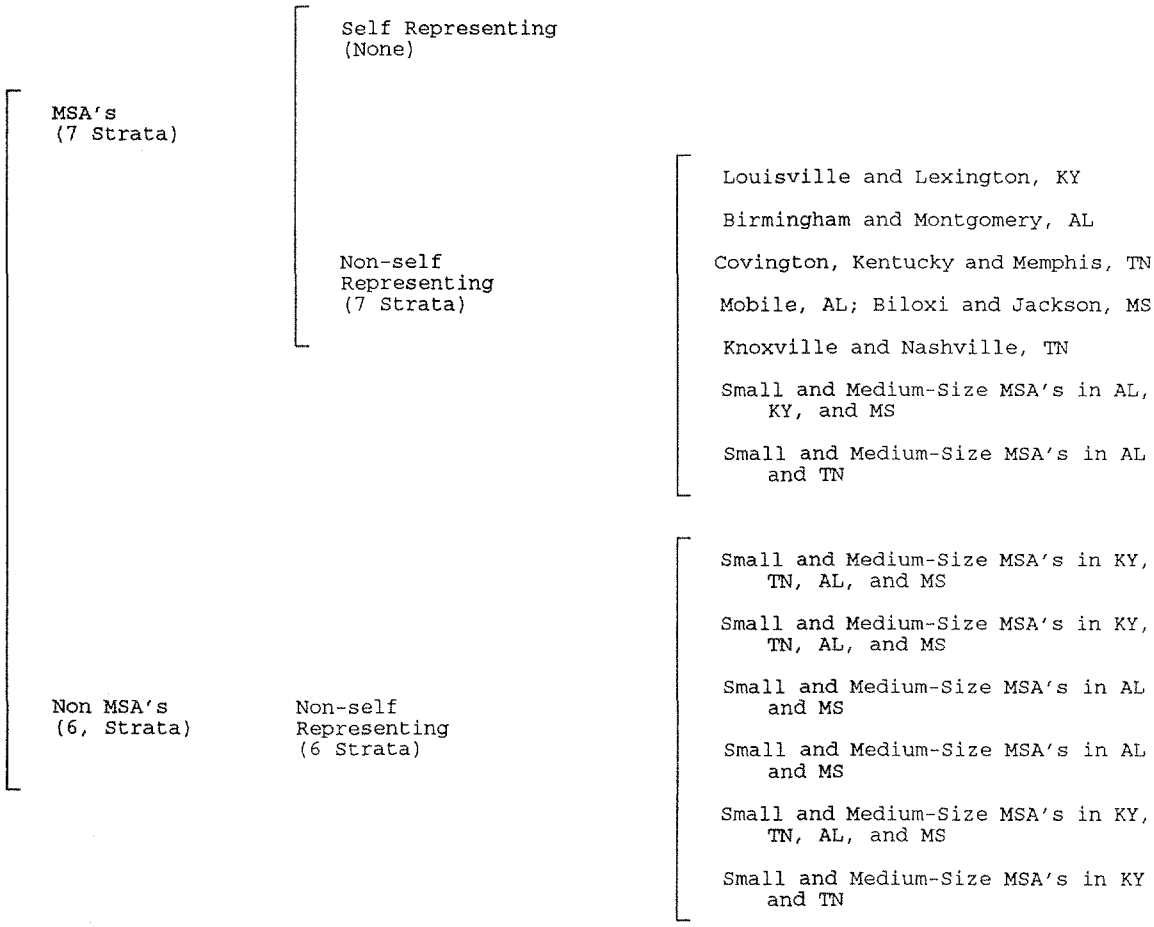
RECS 1980
Intersection 4: Delaware, District of Columbia, Maryland, Virginia, and West Virginia
(6 Strata)



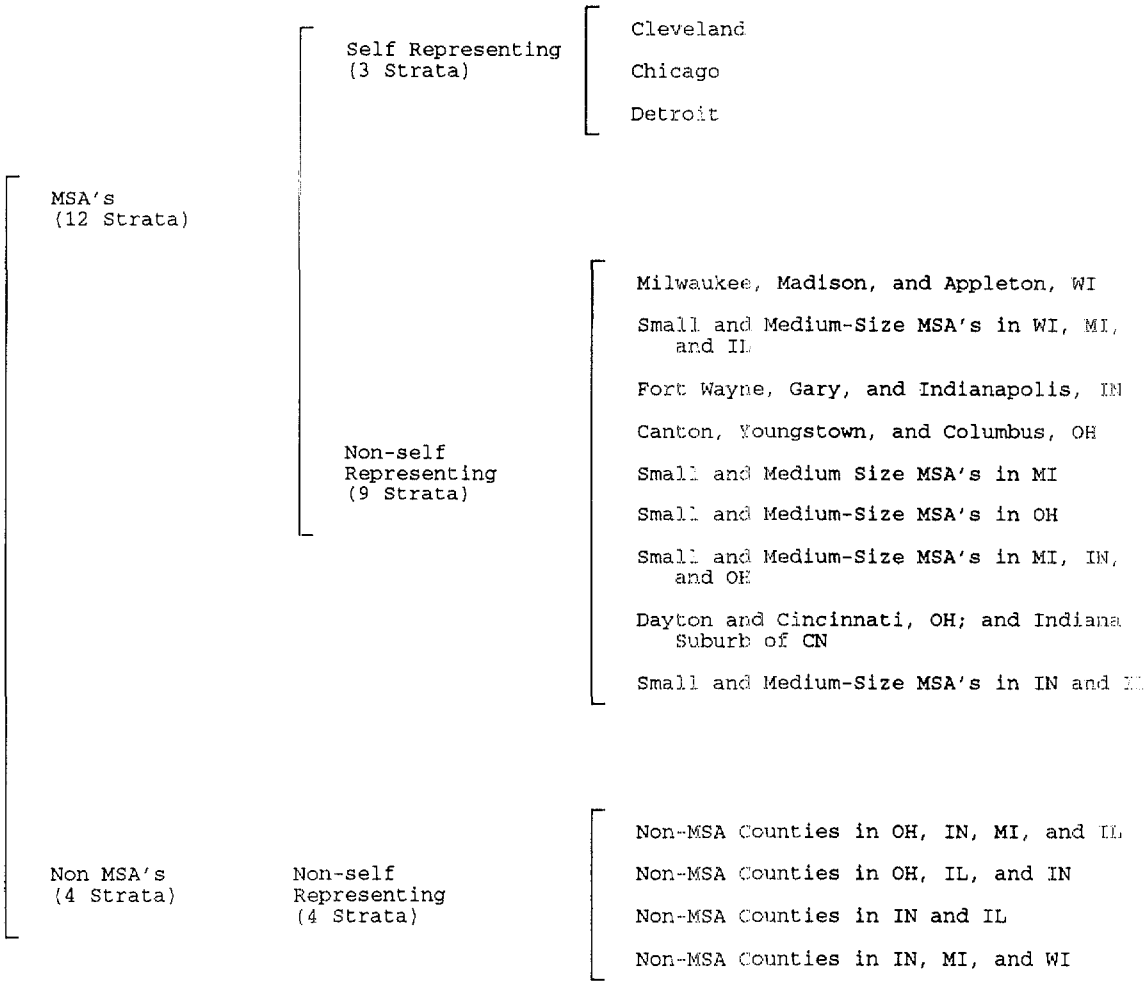
RECS 1980
Intersection 5: Florida, Georgia, North Carolina, and South Carolina
(11 Strata)



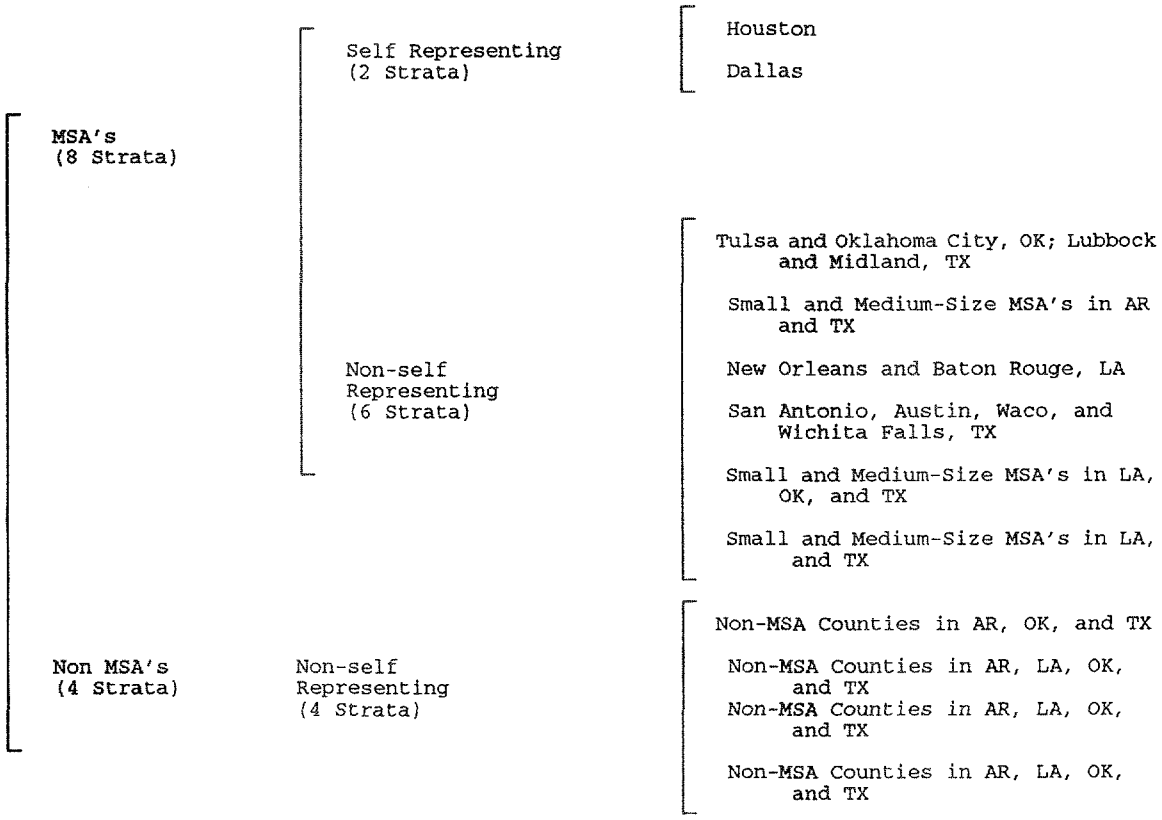
RECS 1980
Intersection 6: Alabama, Kentucky, Mississippi, and Tennessee
(13 Strata)



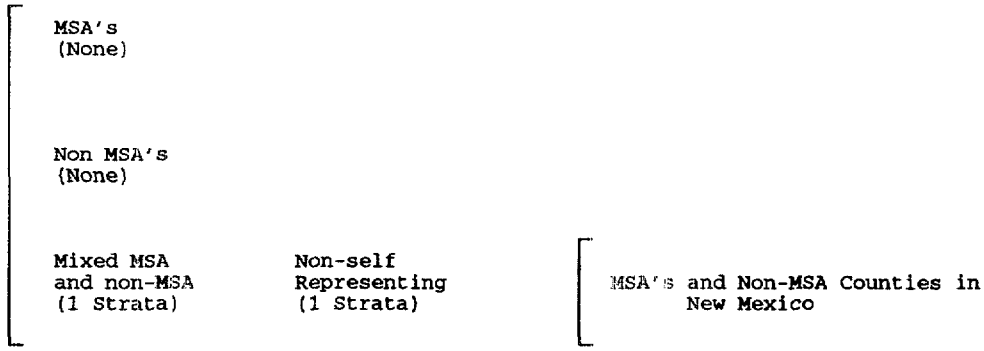
RECS 1980
Intersection 7: Illinois, Indiana, Michigan, Ohio, and Wisconsin
(16 Strata)



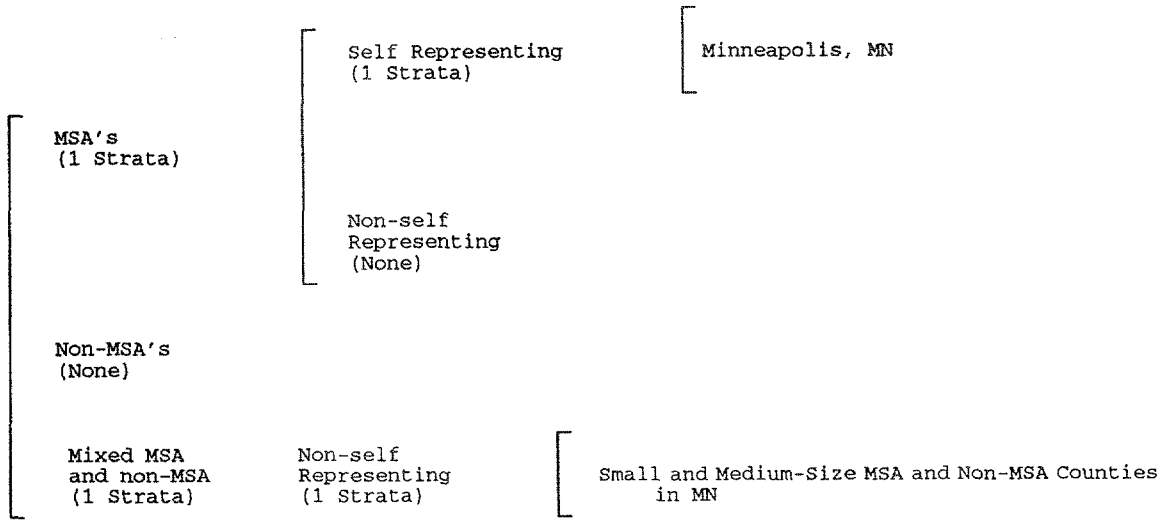
RECS 1980
Intersection 8: Arkansas, Louisiana, Oklahoma, and Texas
(12 Strata)



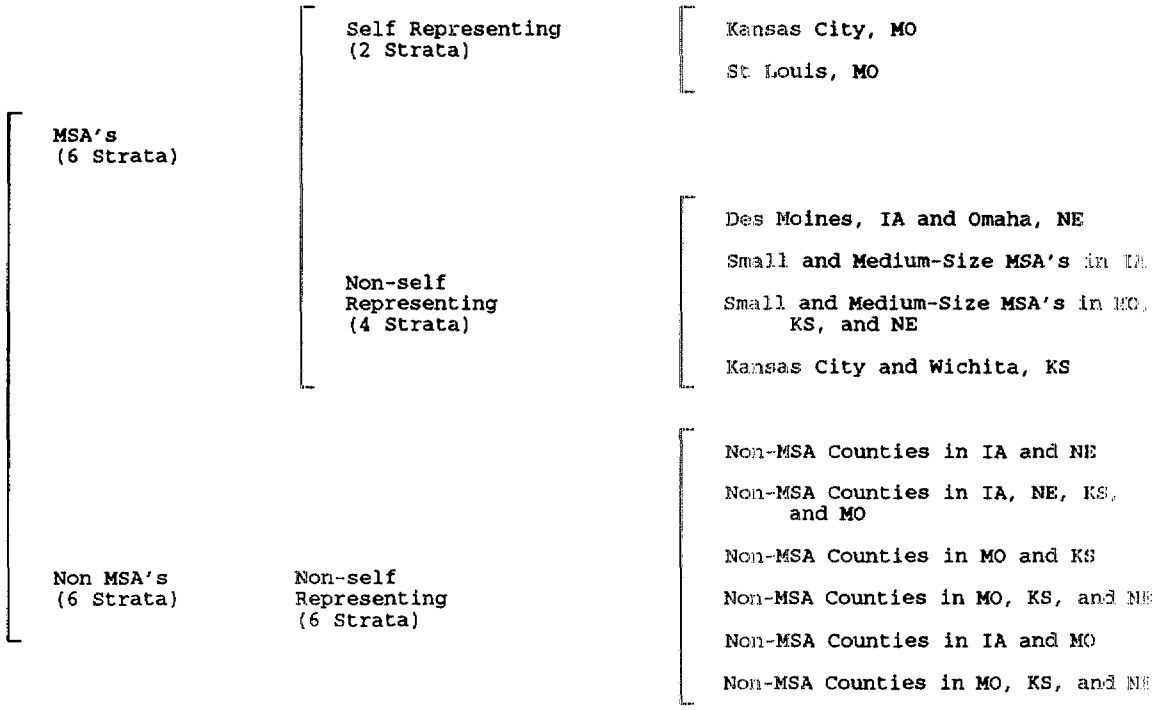
RECS 1980
Intersection 9: New Mexico
(1 Strata)



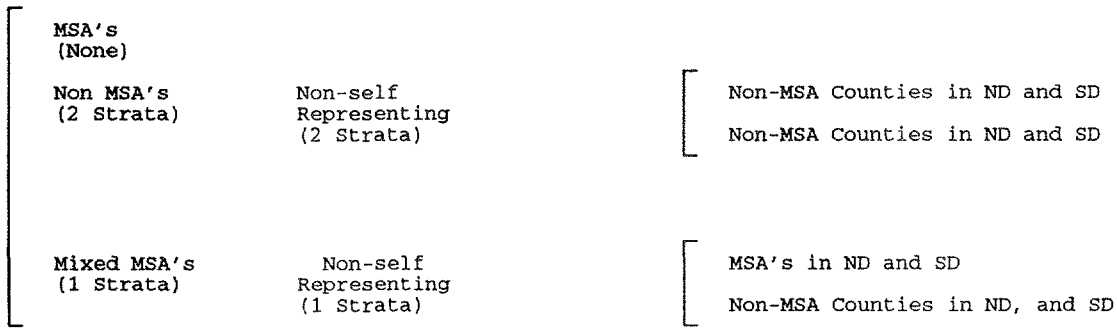
RECS 1980
Intersection 10: Minnesota
(2 Strata)



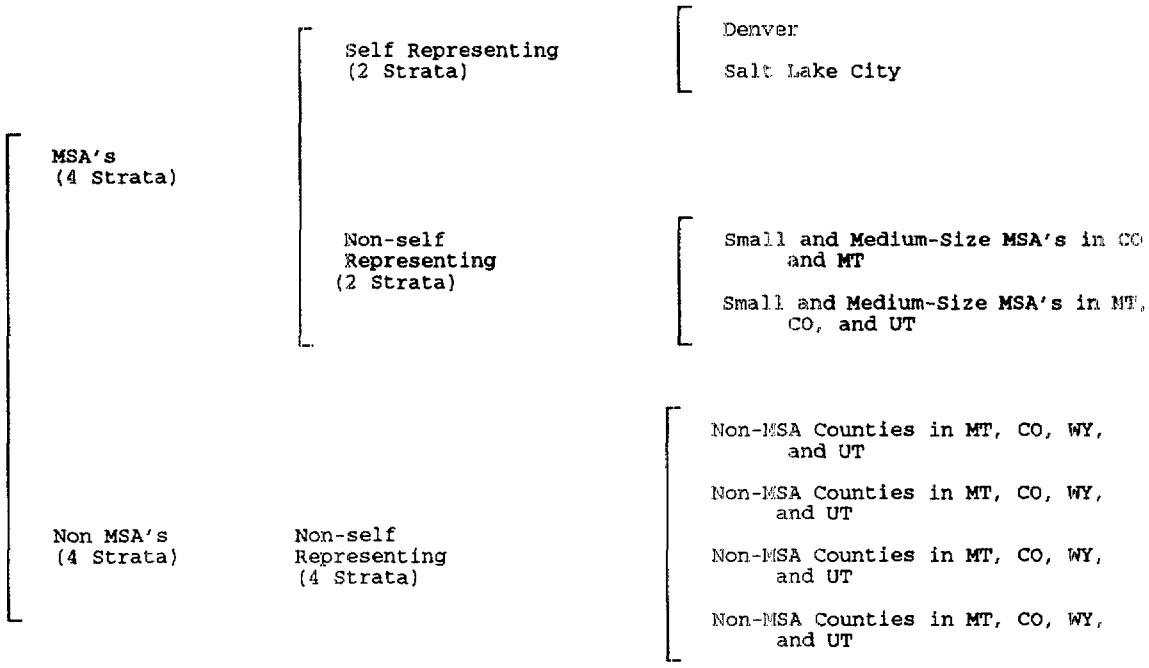
RECS 1980
Intersection 11: Iowa, Kansas, Missouri, and Nebraska
(12 Strata)



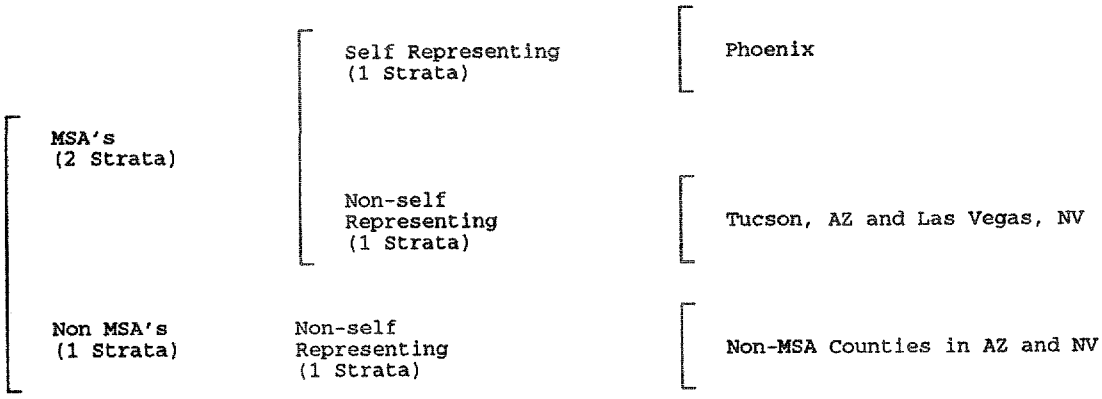
RECS 1980
Intersection 12: North Dakota and South Dakota
(3 Strata)



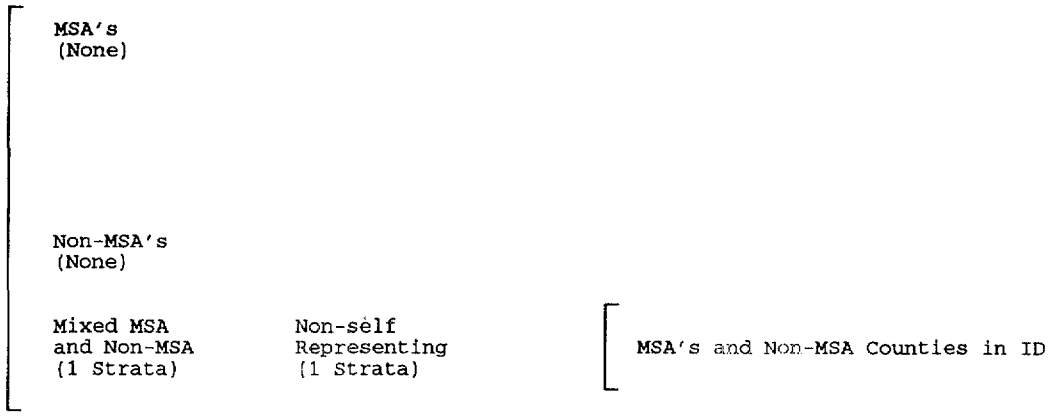
RECS 1980
Intersection 13: Colorado, Montana, Utah, and Wyoming
(8 Strata)



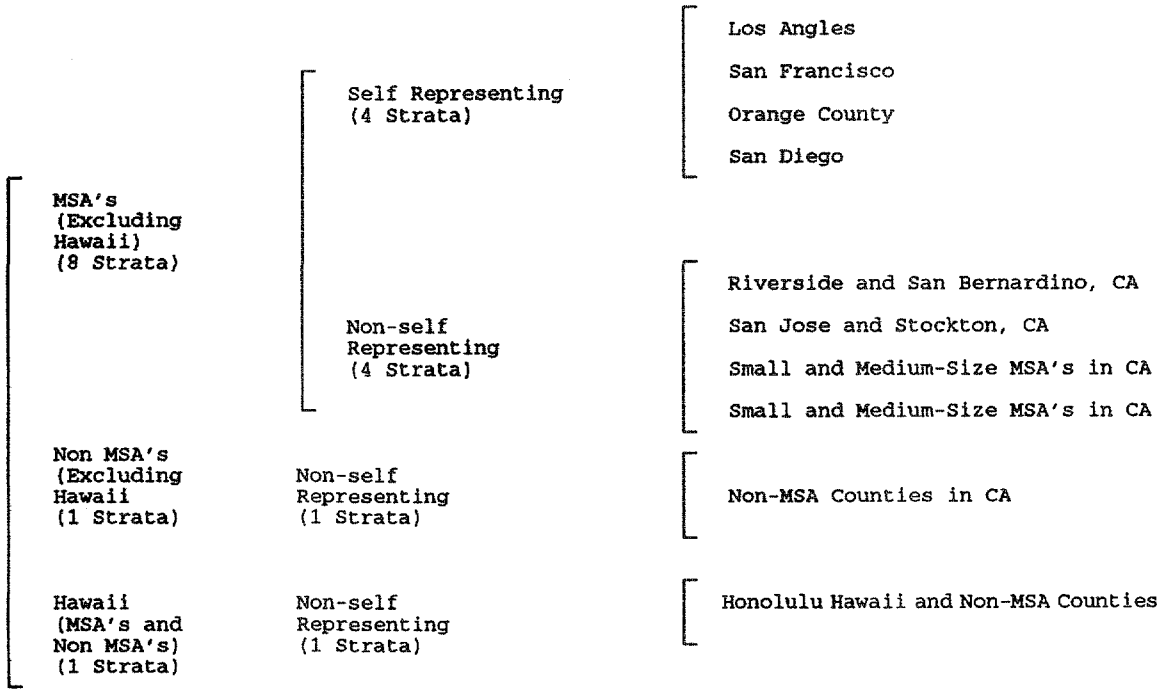
RECS 1980
Intersection 14: Arizona and Nevada
(3 Strata)



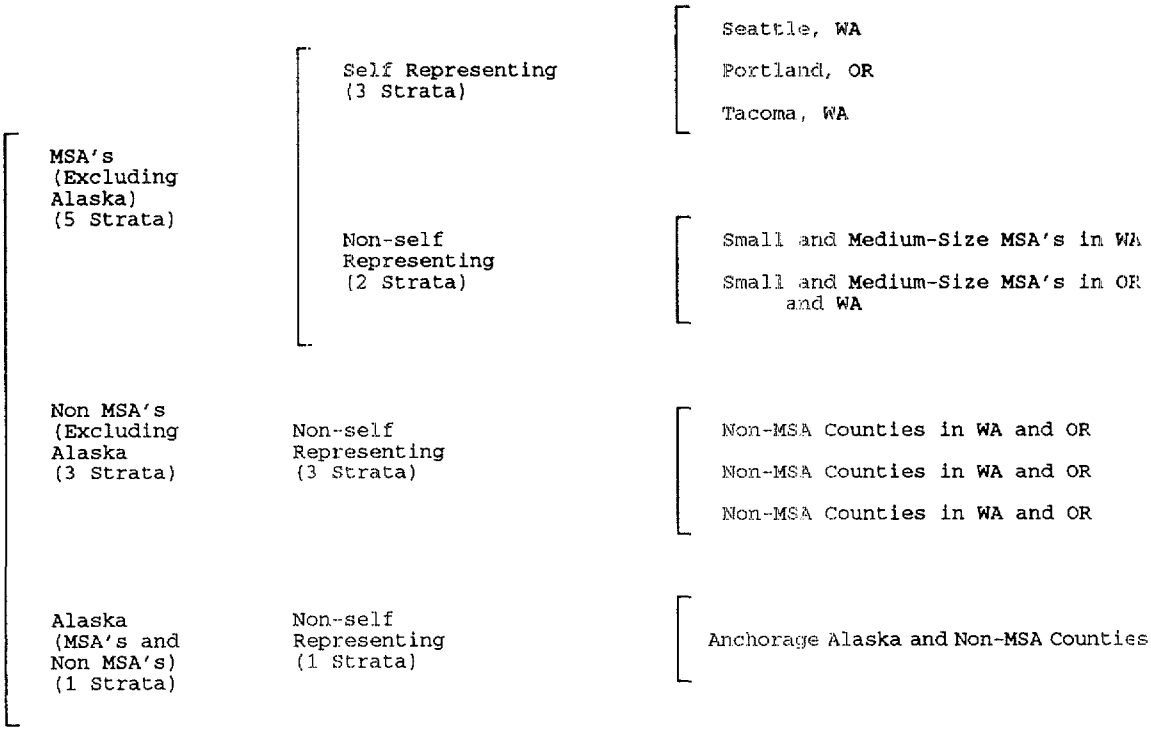
RECS 1980
Intersection 15: Idaho
(1 Strata)



RECS 1980
Intersection 16: California and Hawaii
(10 Strata)

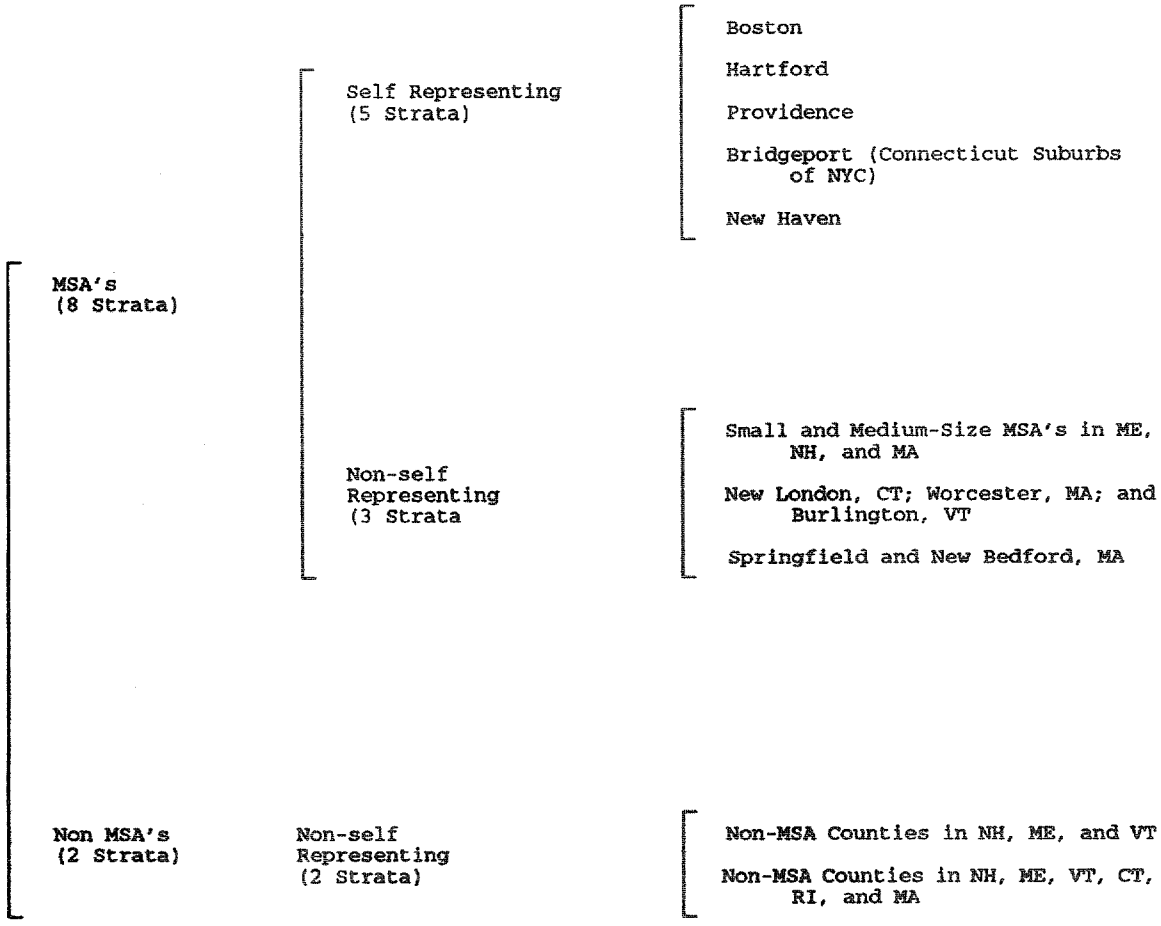


RECS 1980
Intersection 17: Alaska, Oregon, and Washington
(9 Strata)

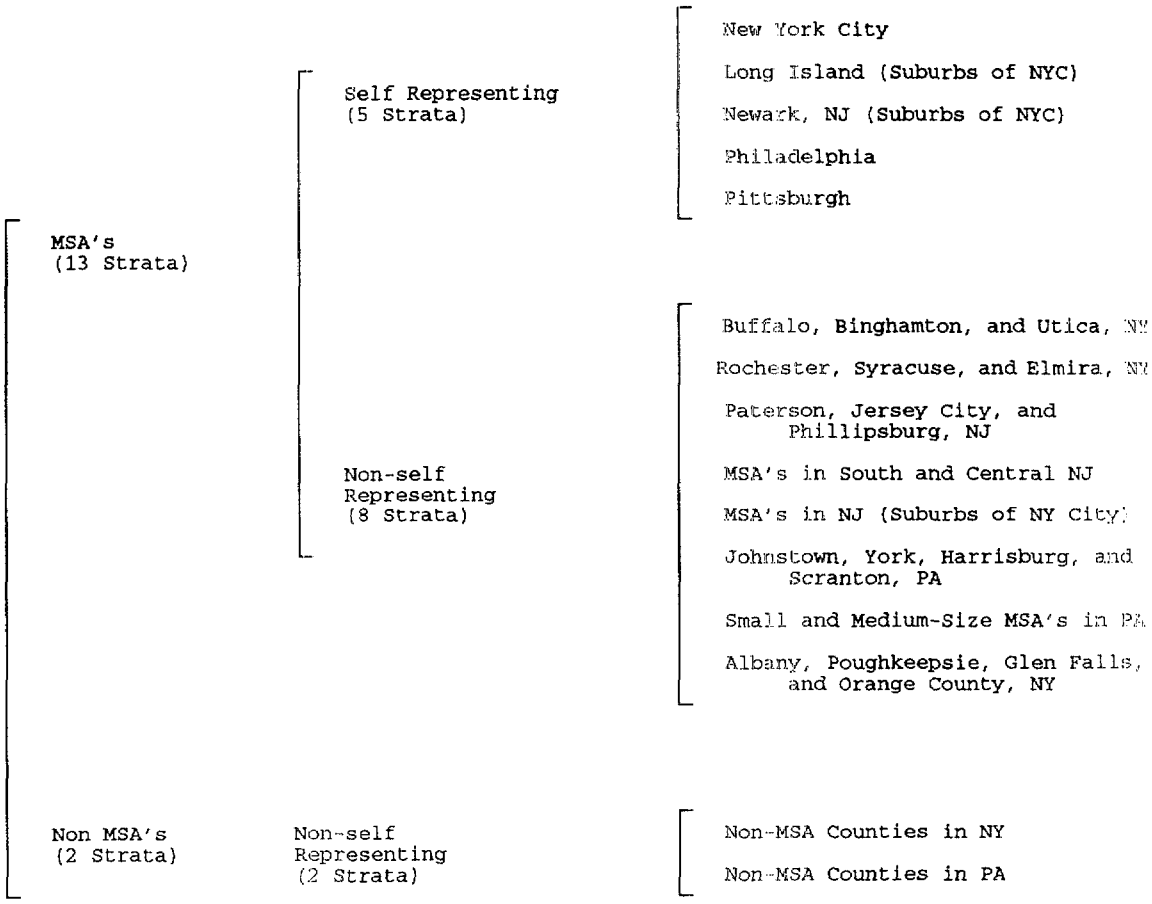


RECS 1984

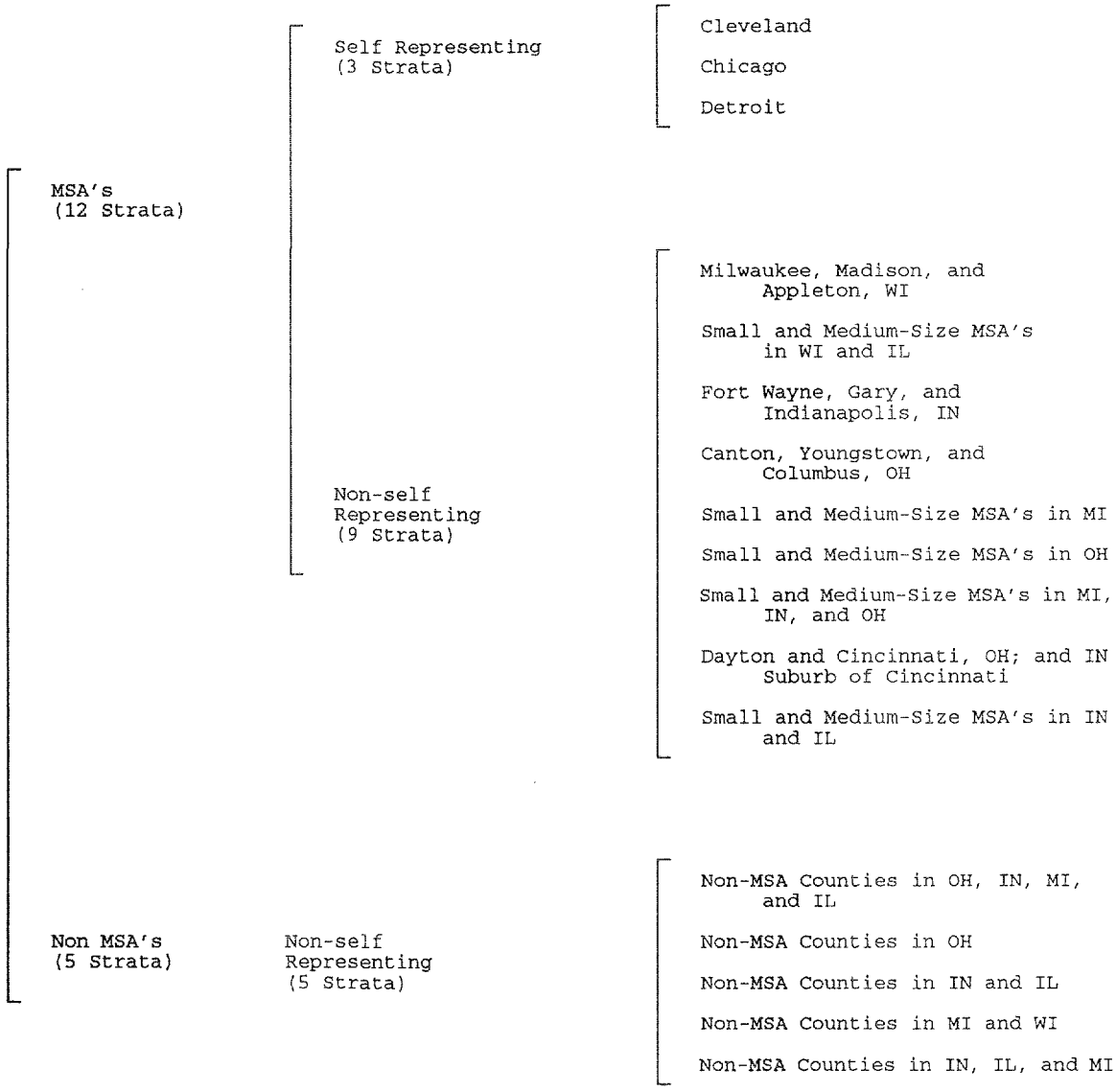
New England Census Division
(10 Strata)



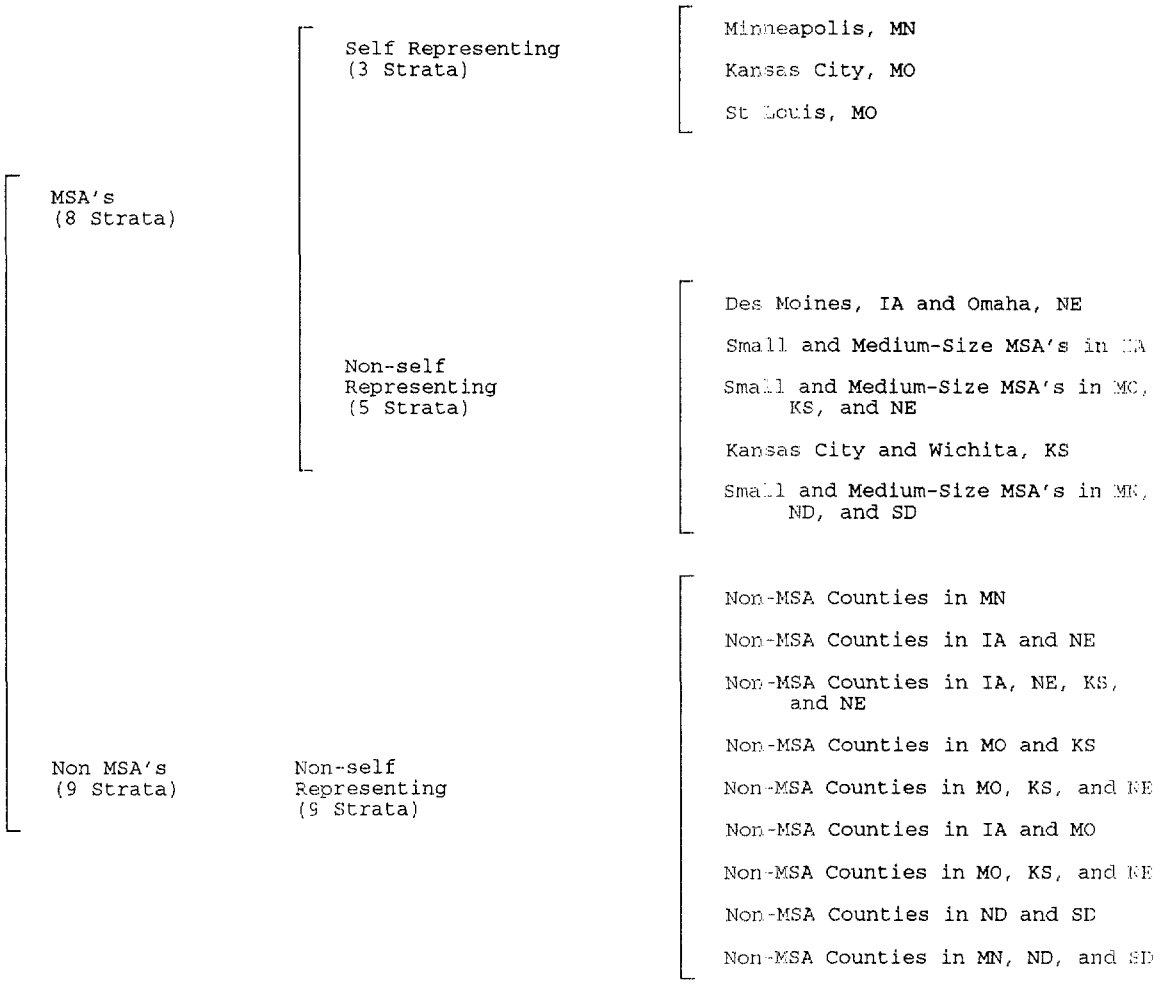
RECS 1984
Middle Atlantic Census Division
(15 Strata)



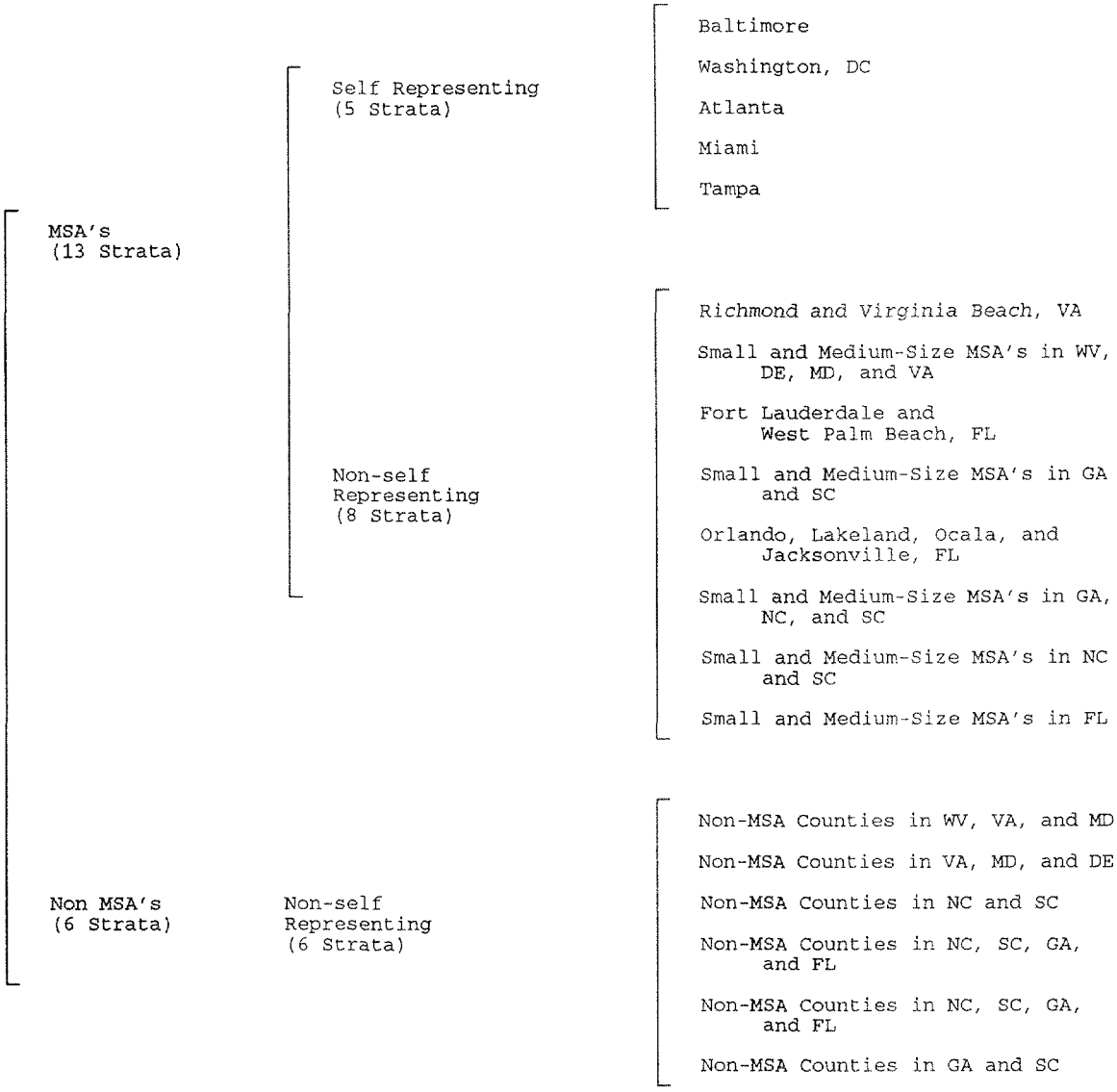
RECS 1984
East North Central Census Division
(17 Strata)



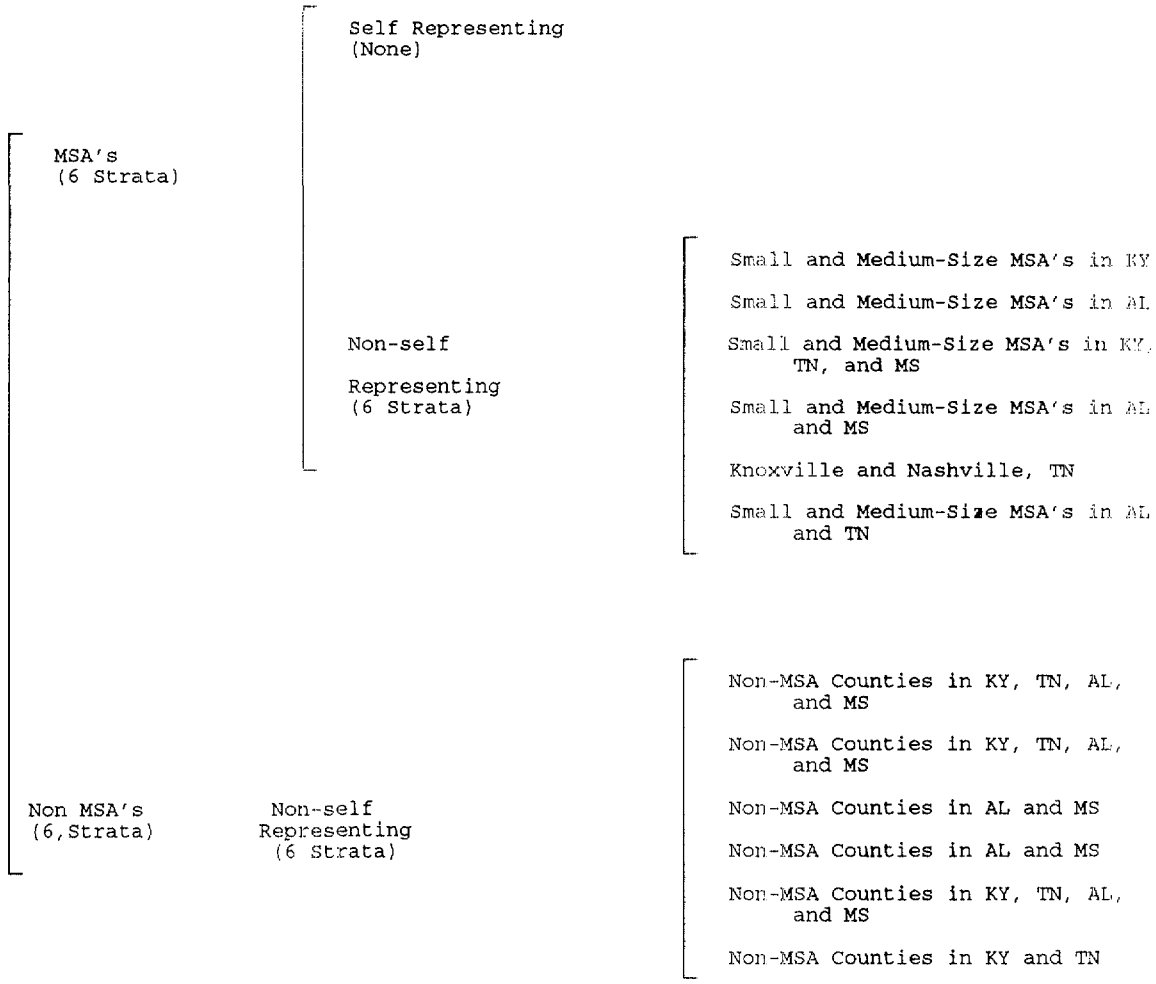
RECS 1984
West North Central Census Division
(17 Strata)



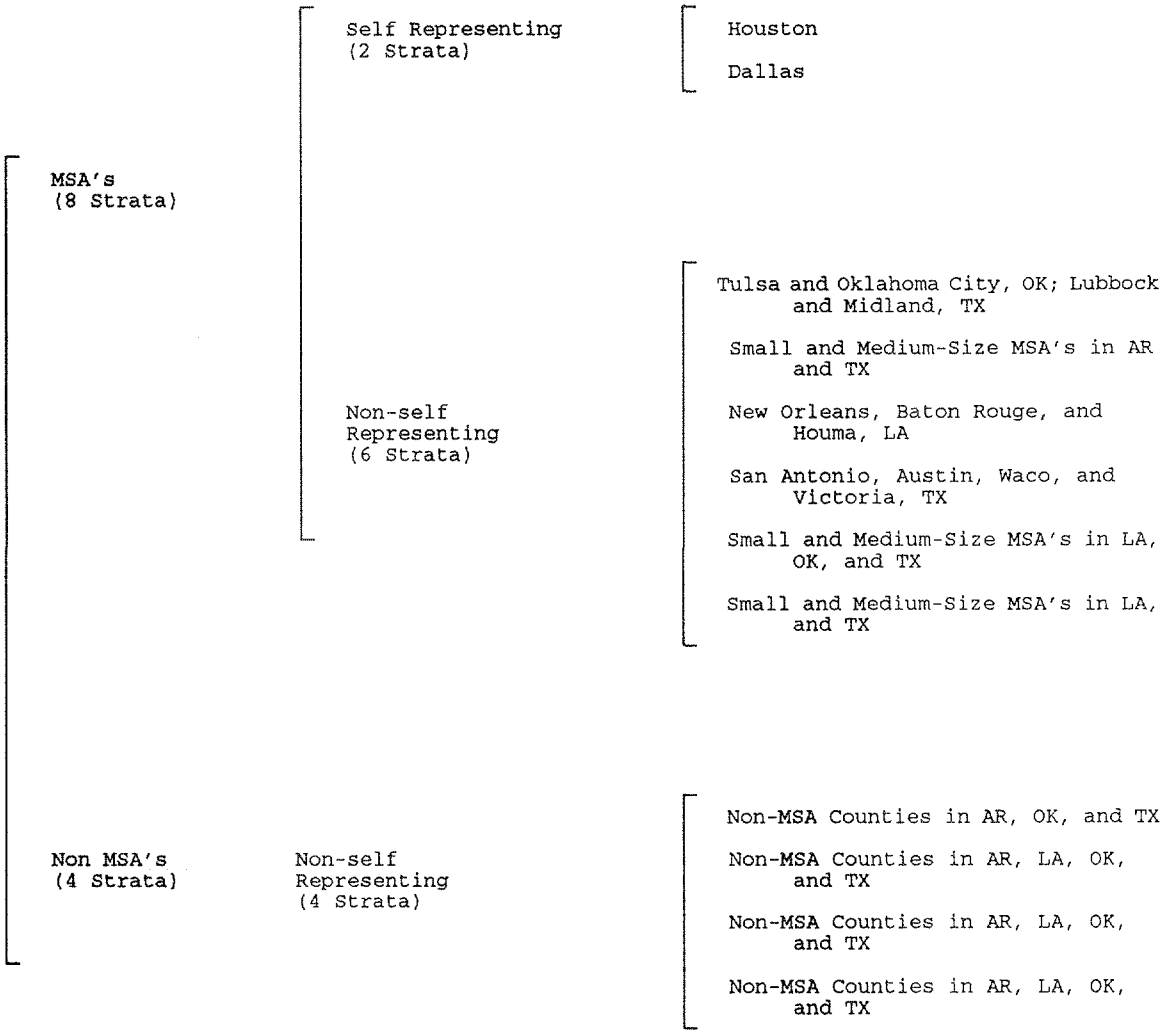
RECS 11984
South Atlantic Census Division
(19 Strata)



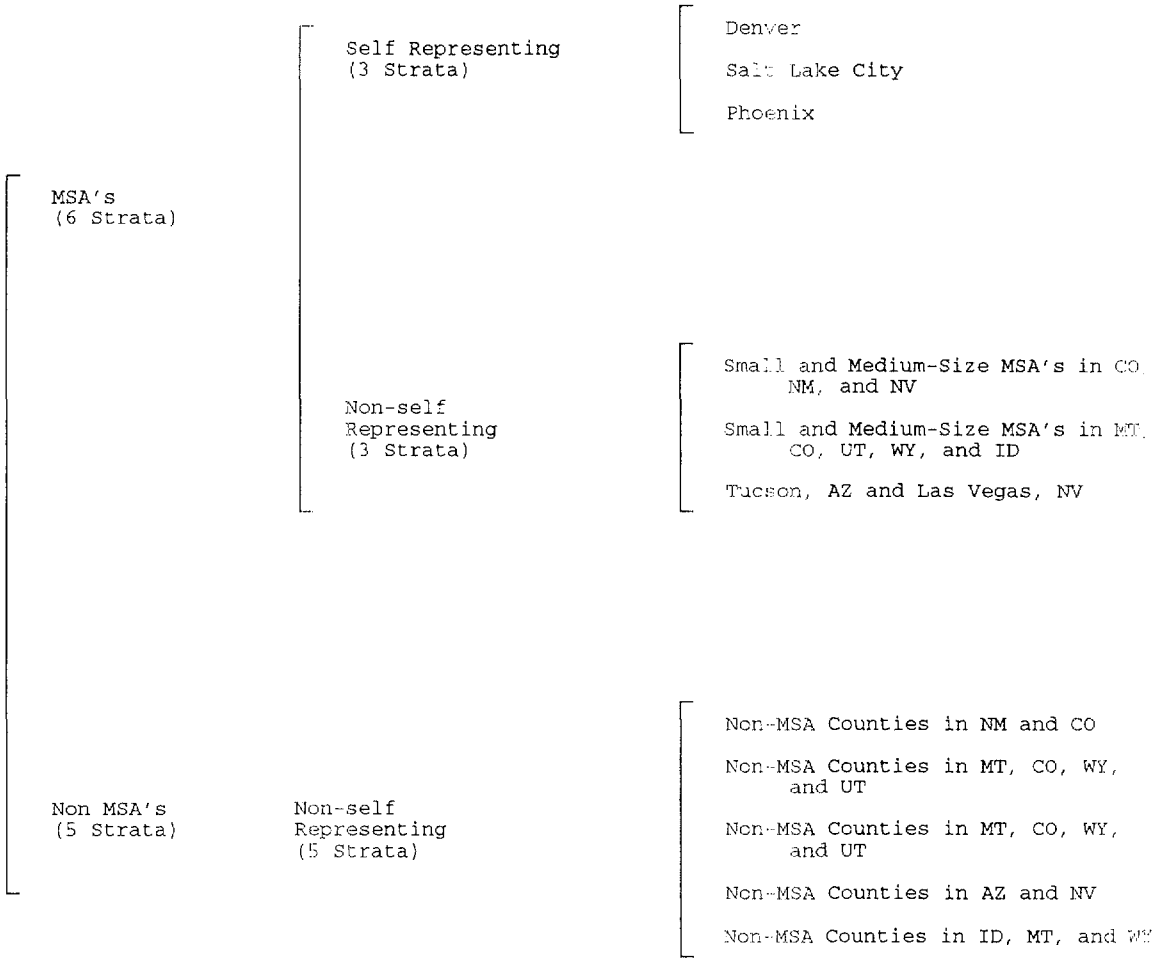
RECS 1984
East South Central Census Division
(12 Strata)



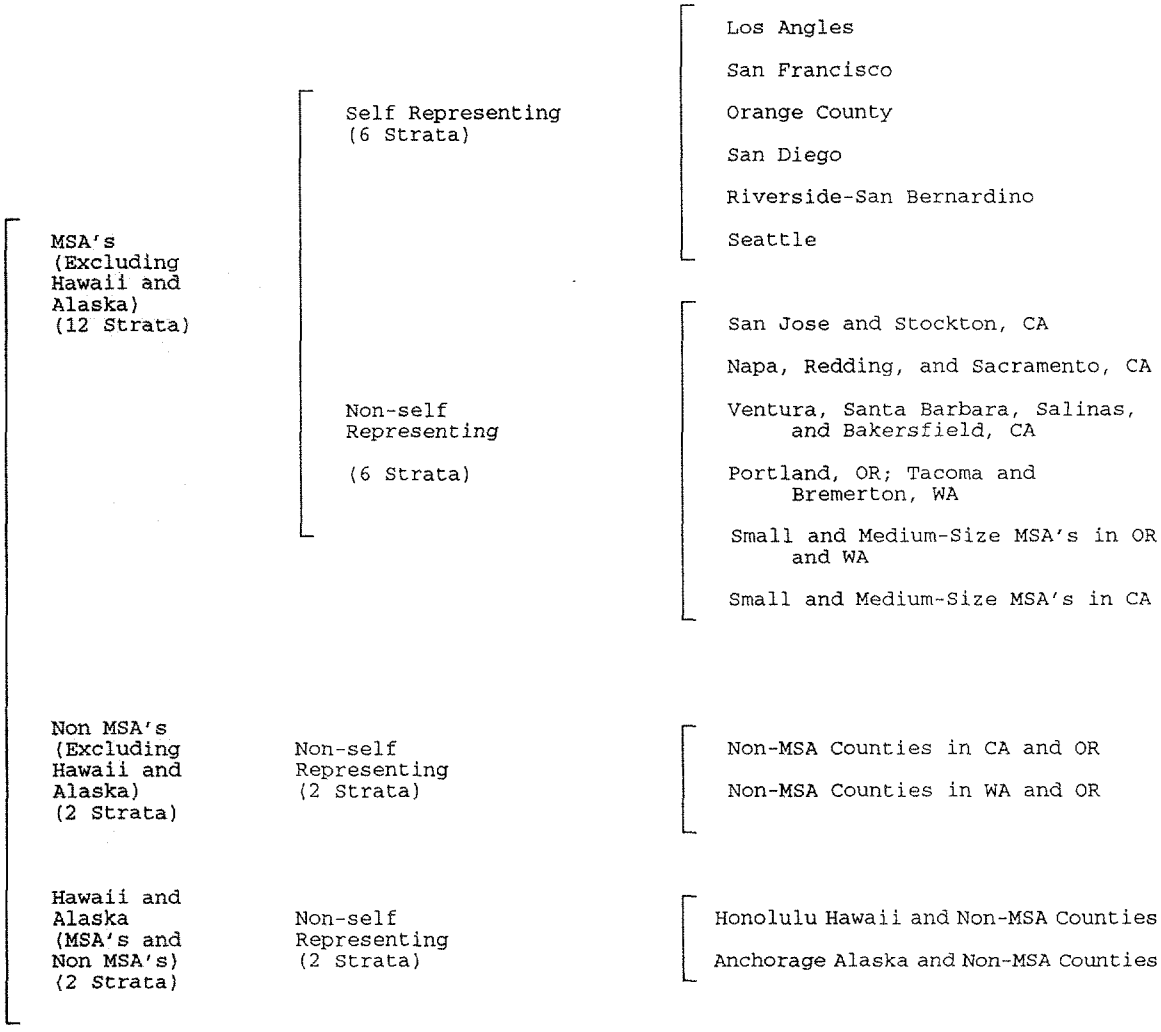
RECS 1984
West South Central Census Division
(12 Strata)



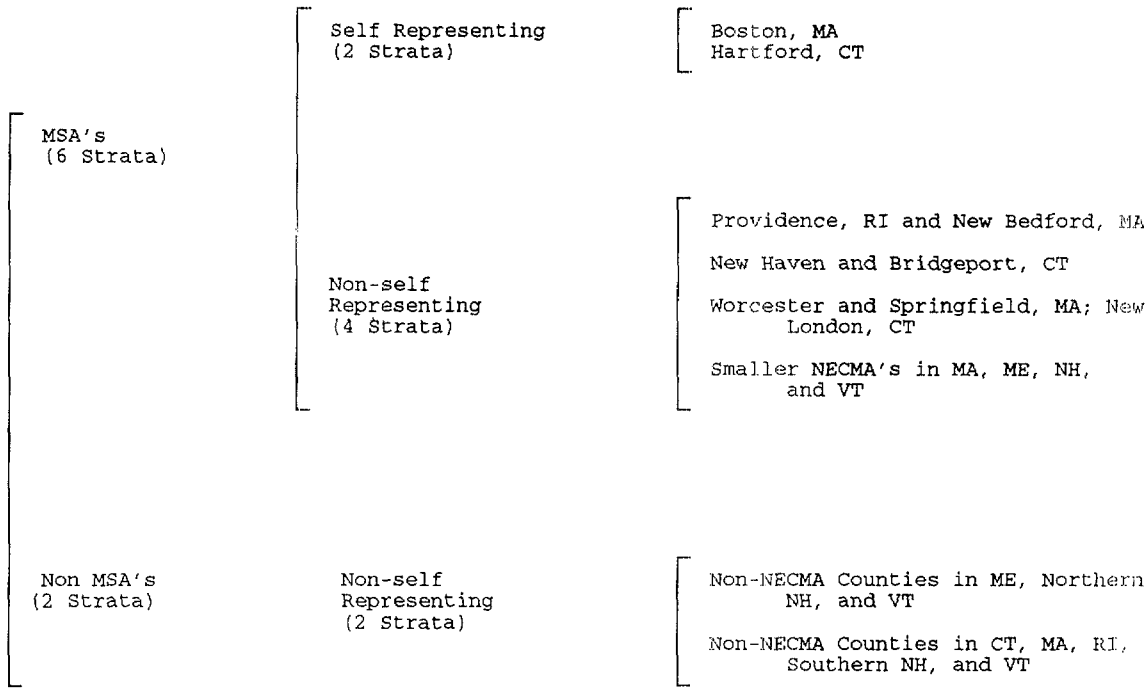
RECS 1984
Mountain Census Division
(11 Strata)



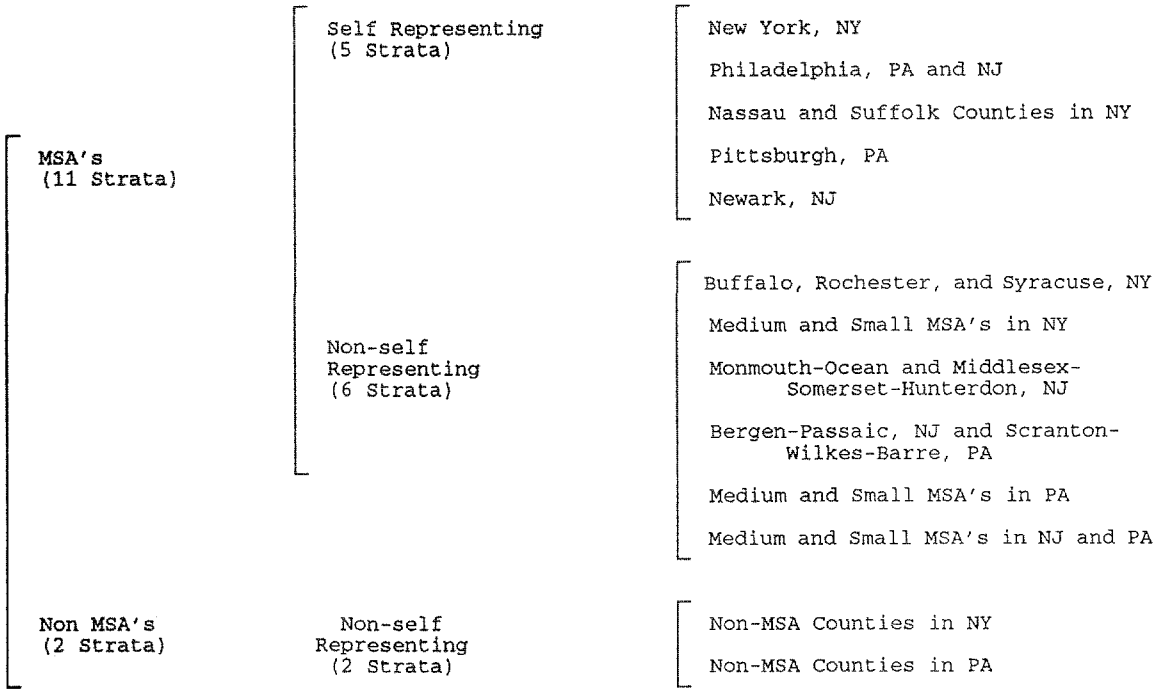
RECS 1984
Pacific Census Division
(16 Strata)



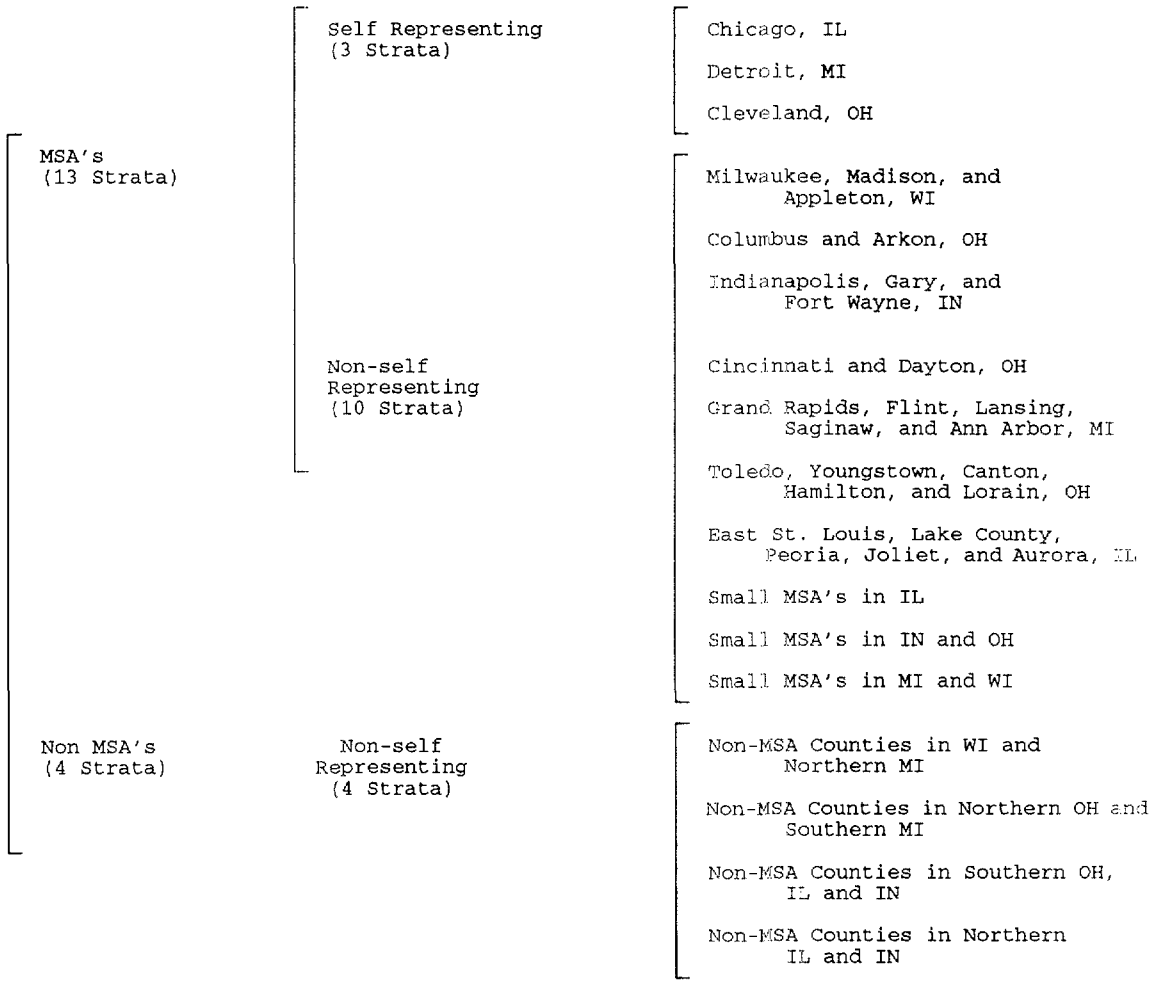
RECS 1993
New England Census Division
(8 Strata – 4,942,714)



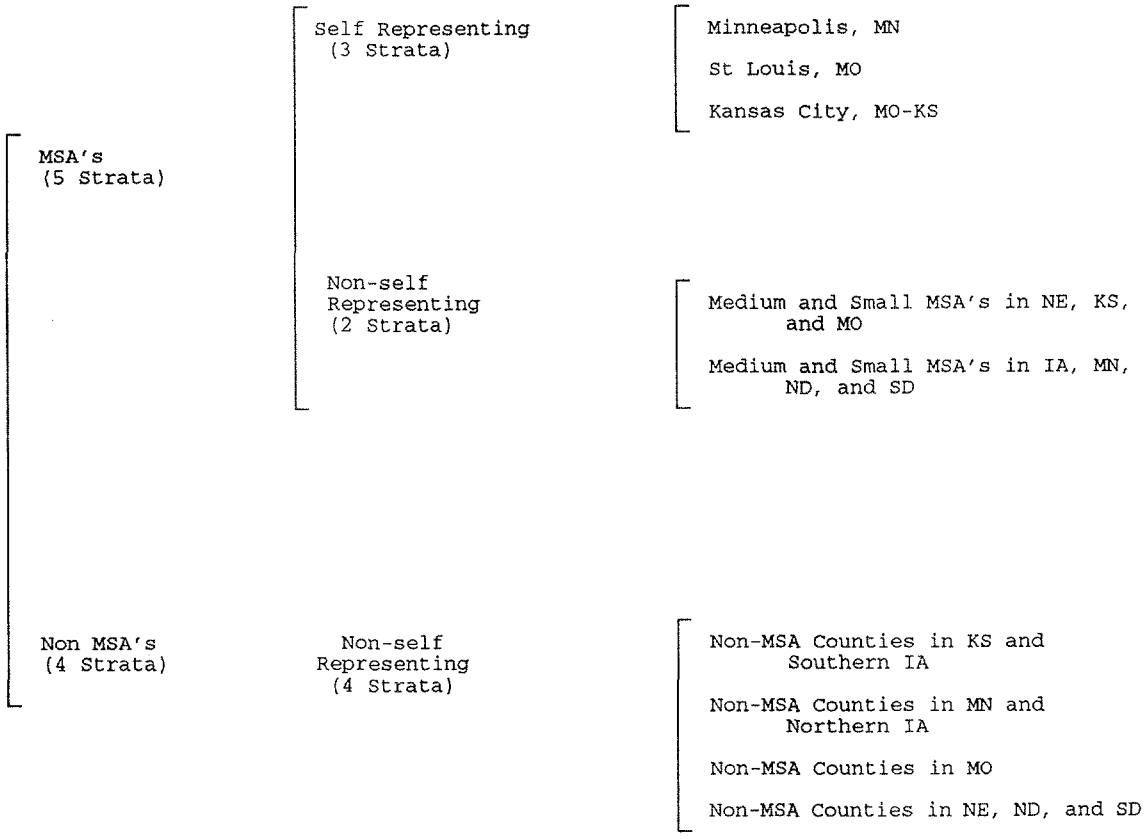
RECS 1993
Middle Atlantic Census Division
(13 Strata)



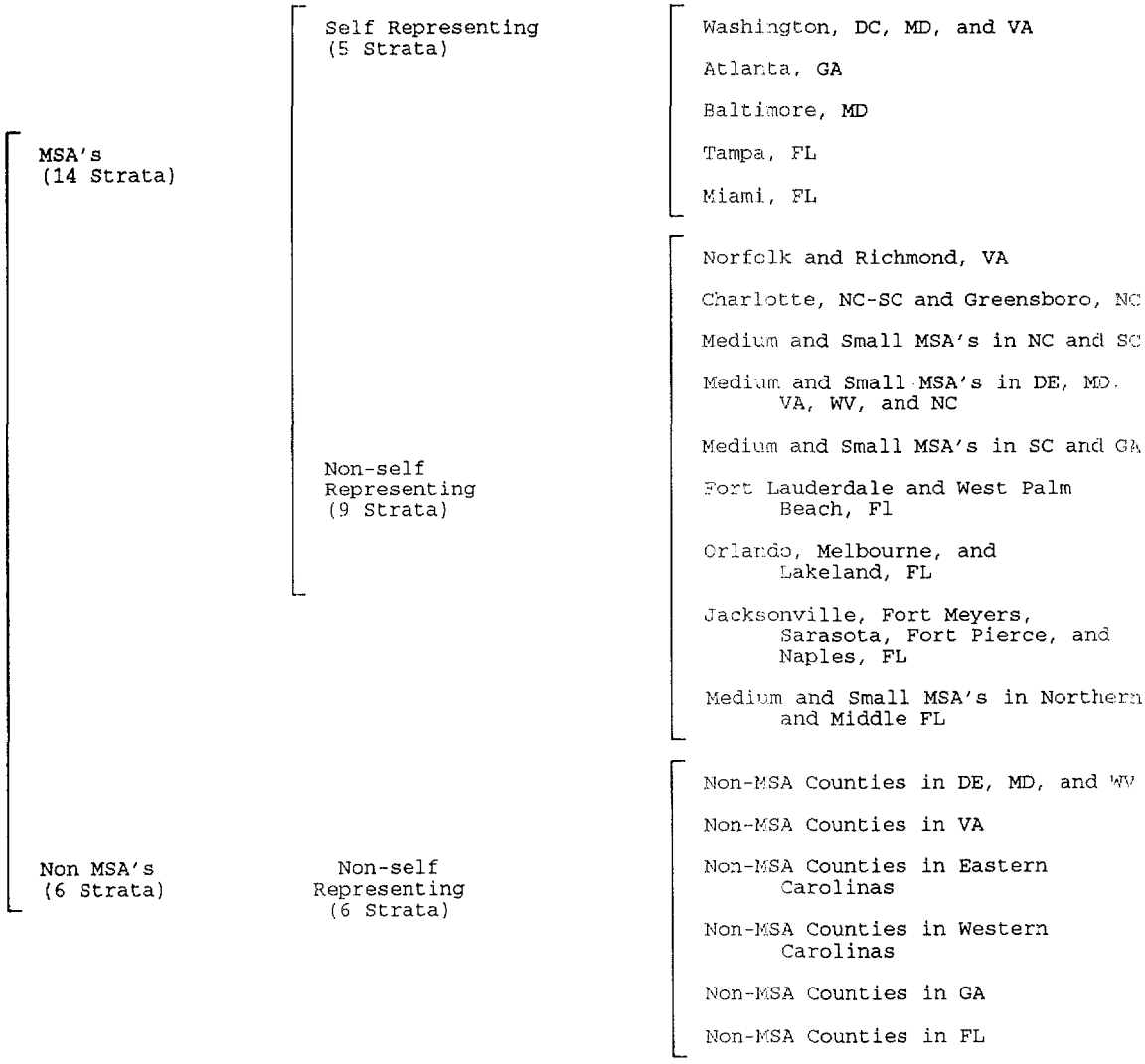
RECS 1993
East North Central Census Division
(17 Strata)



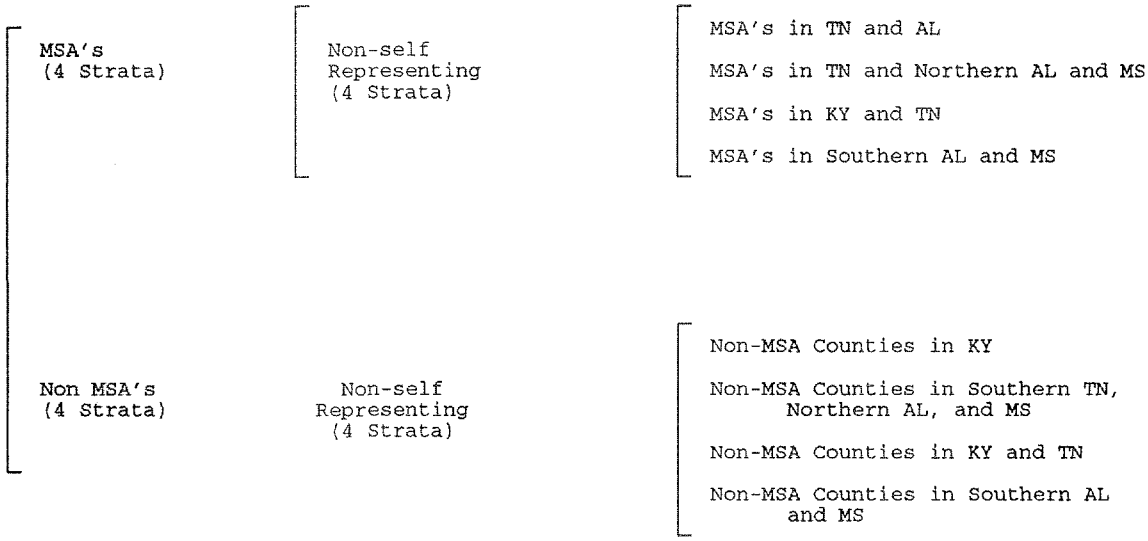
RECS 1993
West North Central Census Division
(9 Strata)



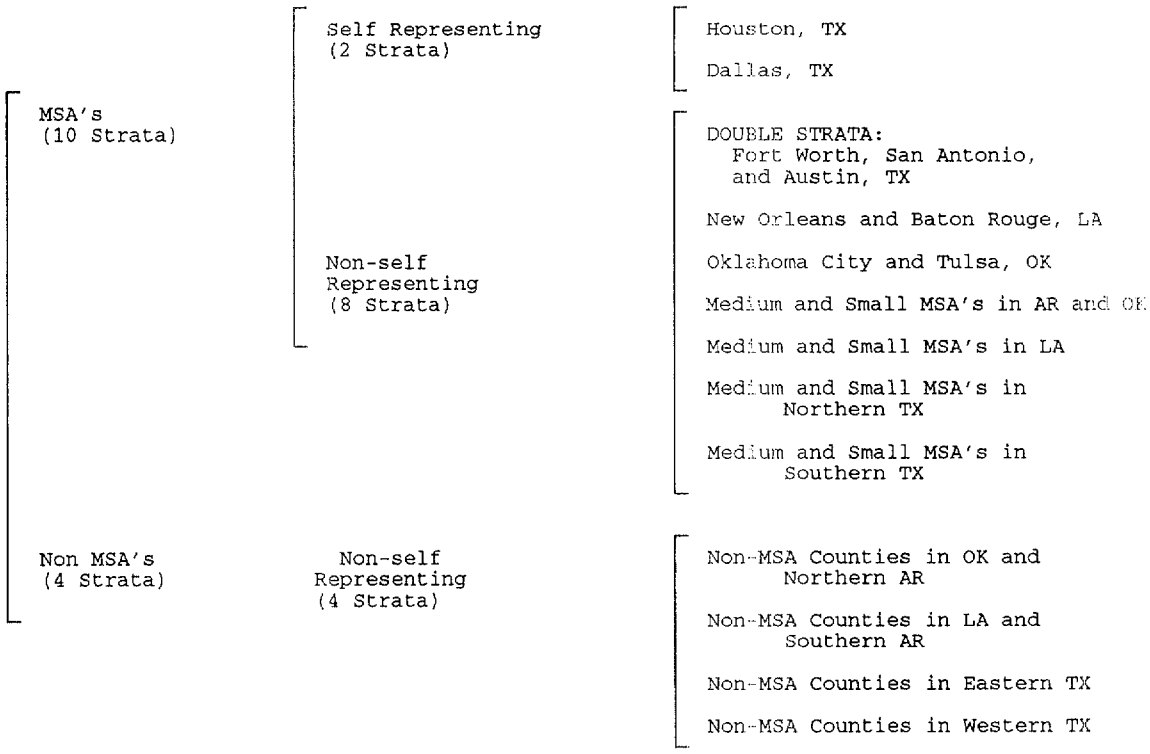
RECS 1993
South Atlantic Census Division
(20 Strata)



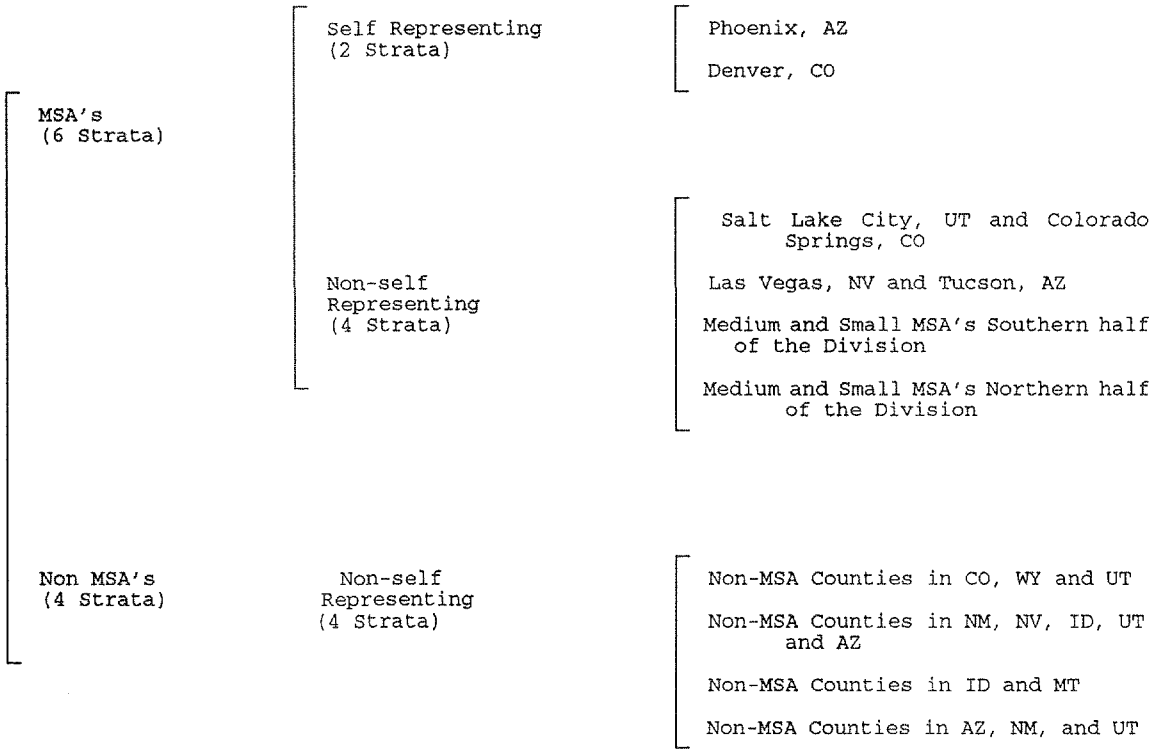
RECS 1993
East South Central Census Division
(8 Strata)



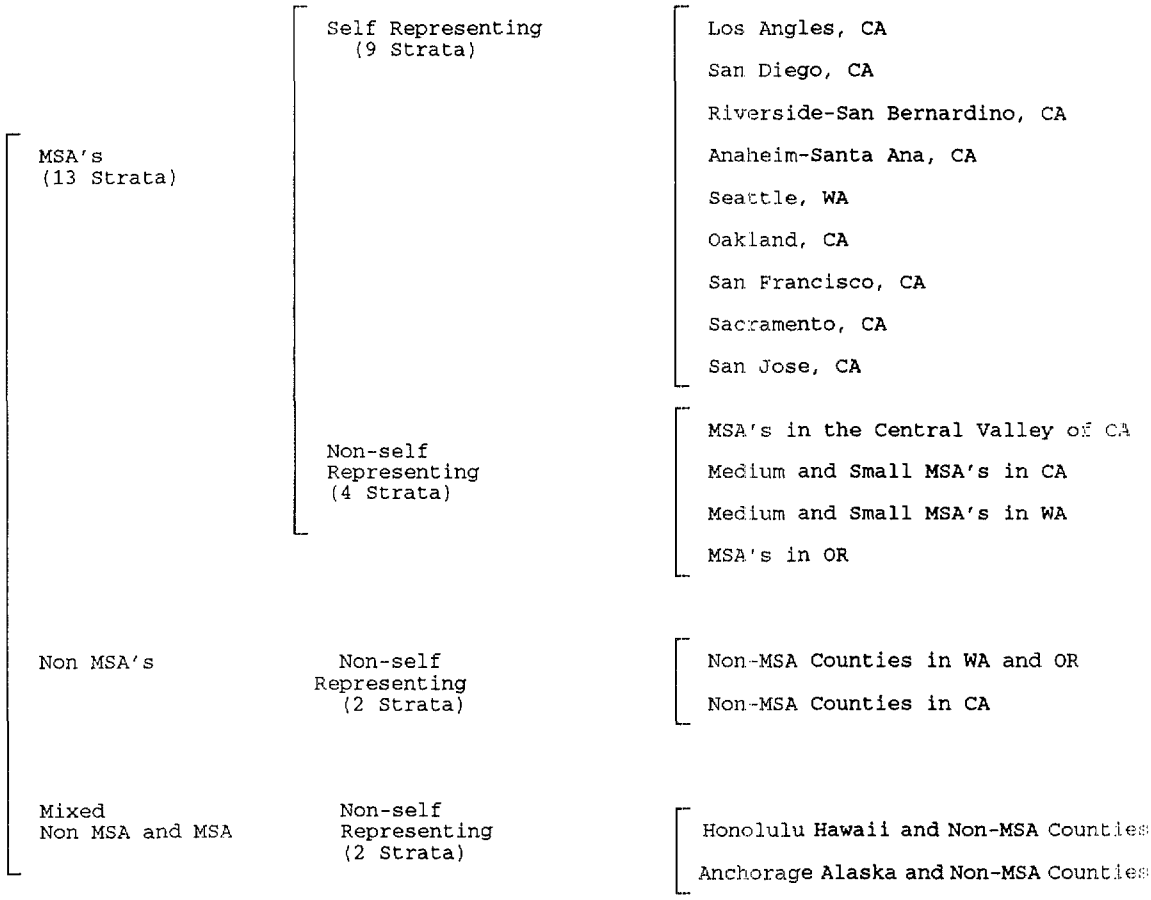
RECS 1993
West South Central Census Division
(14 Strata)



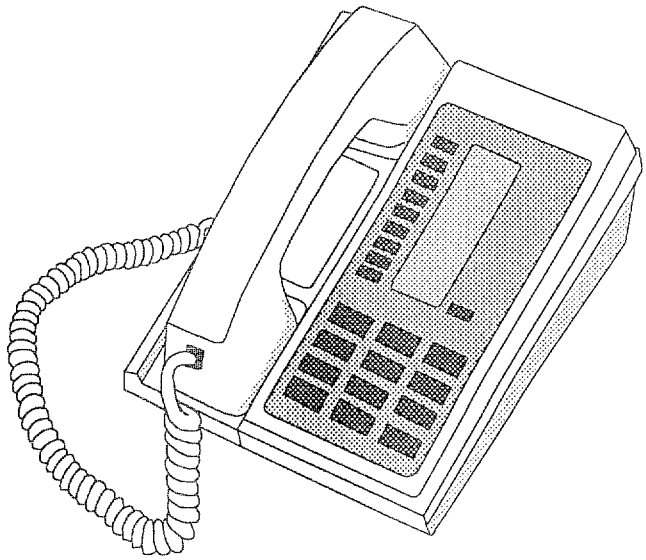
RECS 1993
Mountain Census Division
(10 Strata)



RECS 1993
Pacific Census Division
(17 Strata)



Call for your copy of *EIA New Releases*



A concise and timely review of new statistical publications and energy news

Offered from the National Energy Information Center

For a complimentary copy of this bimonthly bulletin, or a complimentary subscription, call (202)586-8800.

Energy Information Administration
U.S. Department of Energy
Forrestal Building, EI-231
Washington, DC 20585

THIRD-CLASS MAIL
POSTAGE & FEES PAID
U.S. DEPT OF ENERGY
PERMIT NO. G-[illegible]

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

Address Correction Requested