

THE ELCAP RESIDENTIAL BASE STUDY SAMPLE:  
A BASIC CHARACTERIZATION

Phillip A. Windell  
Bonneville Power Administration

ABSTRACT

The End-use Load and Conservation Assessment Program (ELCAP) sponsored by the Bonneville Power Administration is among the first large scale fuel monitoring projects. This paper is a product of initial analytical work directed at assessing the results of the residential sampling and recruiting efforts, and is intended as an early contribution to the development of large scale monitoring protocols. Following a brief overview of the ELCAP and a summary description of the residential sample selection process, the paper then presents an initial characterization of the final residential Base Study sample. The paper concludes with an assessment of the relative success of the residential sample selection strategy in achieving the original objectives for the study.

THE ELCAP RESIDENTIAL BASE STUDY SAMPLE:  
A BASIC CHARACTERIZATION

Phillip A. Windell, Bonneville Power Administration

INTRODUCTION

Recent advances in sensor and micro-electronics technologies have lead to substantial reductions in the costs associated with field monitoring of energy usage, especially electricity. It is now becoming economically feasible to conduct relatively large scale, high resolution energy monitoring studies. As a result, rather than suffering from a dirth of even low quality intermittent data, energy analysts of the future are likely to have access to a large volume of relatively high quality data. The character and pace of the intervening growth period will depend on the extent to which experiences are shared in the open literature and absorbed by those intending to initiate large scale monitoring projects. By describing critical portions of the implementation process and attempting to evaluate the consequences of the strategies employed for an early large scale monitoring program, this paper seeks to contribute to this learning process.

Even a traditional interview-based survey is a costly affair, however. Since this is one of the earliest large scale monitoring studies, the results are likely to be of interest to a rather large and varied audience, including significant numbers of analysts, planners, and technicians outside the particular geographic area in which the data are being collected. Prior to accessing the data, however, analysts are likely to be interested in gaining some basic understanding of the distributions of certain fundamental characteristics of the participating sites. This is a second purpose of the paper: to serve as a point of departure for further analyses.

PROGRAM OVERVIEW

The End-use Load and Conservation Assessment Program (ELCAP) is an end-use electricity monitoring program sponsored by the Bonneville Power Administration (Bonneville) to support its resource management responsibilities. The ELCAP was initiated in August 1983 under an intra-agency agreement with the Pacific Northwest Laboratory operated for the U.S. Department of Energy by Battelle Memorial Institute. Under the ELCAP, approximately 600 residential, commercial and multifamily structures are being instrumented with data logging equipment (Tomich and Schuster, 1985). In addition to the energy use data, the loggers record at least one indoor temperature and 70 sites have been outfitted with a complete set of outdoor meteorological sensors. Although the integration period is variable and may be set as low as a minute, as a standard, data are recorded at hourly intervals. The data are acquired from each of the sites approximately every two weeks over standard telephone lines, routinely verified, and placed in a central data repository. A full set of commercially available data management and computational tools are being integrated in a custom software environment to support ongoing analysis activities<sup>1</sup>. In addition to the energy and meteorological data, the database includes a full array of building, occupant, and data status characterizations.

This paper is concerned exclusively with the residential portion of the ELCAP, and particularly with the Base Study. Specifically, the paper includes a summary description of the procedures used to select the sample, provides a summary characterization of the final sample, and evaluates the final sample relative to several criteria of 'success.'

## SAMPLE SELECTION<sup>2</sup>

### PNWRES-83 - The Sample Frame

In 1983, the Bonneville Power Administration sponsored a regional survey of residential energy consumption. The 1983 Pacific Northwest Residential Energy Survey (PNWRES-83) involved personal interviews with a probability sample of 4,703 households selected from electric utility metering accounts (L. Harris & Associates, 1983, 1984a, 1984b). The resulting database constituted a powerful frame for the selection of the ELCAP residential sample.

The selection of the sample was constrained by a number of parameters, not the least of which was budgetary. Based on initial cost estimates and the available budget, the target for the Residential Base Study was fixed at 400 units. While this is quite large by end-use monitoring standards, it is extremely modest by survey research standards. For the results to be useful, therefore, it was clearly necessary to treat the study more as an experiment than a traditional sample survey. That is to say, the study would need to focus on a segment of the population and employ hypothesis testing procedures to support inferences and generalizations. As a result of a series of meetings and careful review of the PNWRES-83 data, it was agreed to focus the study on residences which use electricity for space heating purposes. It was further agreed to restrict the Base Study to site-constructed, detached, owner-occupied dwellings. Based on the PNWRES-83 data, this group accounts for approximately 25 percent of the PNW residences (see Figure 1).

In an effort to test the importance of each of the restrictions, four Case Studies were defined. Each Case Study varies one and only one of the major characteristics defining the Base Study sample. Thus, Case Study #1 involves homes which are attached rather than detached, but are otherwise similar to Base Study homes, i.e., site built, owner-occupied, single-family units with permanent electric space heating equipment. Case Study #2 includes renter occupied, as opposed to owner-occupied homes which are otherwise similar to those participating in the Base Study. Case Study #3 involves manufactured as

-----  
<sup>1</sup>An early description of portions of the system is contained in Stokes et al., 1985. A complete description and associated user guides are scheduled for publication in early 1987.

<sup>2</sup>For a complete description of the residential sample design, see Windell, forthcoming.

opposed to site-built homes. Finally, Case Study #4 involves homes which are heated with forced air gas or oil furnaces. Multifamily units are reserved for a separate study under the ELCAP. Due to the small size of the samples for each of the Case Studies, and the space limitations here, the Case Studies are not considered further in this paper.

### Selection of the Base Study Sample

The strategy adopted for the selection of the Base Study sample was consistent with the quasi-experimental, hypothesis testing approach mentioned previously. That is, in order to ensure representation over the full range of relevant characteristics, the frame of eligible dwellings was intensively stratified using the data from the PNWRES-83. Five criteria were employed, as follows:

1. Geographic area
  - a. Western Washington and Oregon
  - b. Eastern Washington and Oregon
  - c. Idaho and Montana
2. Serving electric utility ownership type
  - a. Investor owned (IOU)
  - b. Publicly owned (POU)
3. Year dwelling constructed
  - a. Before 1960
  - b. 1960 - 1978
  - c. After 1978
  - d. Missing
4. 1982 household income
  - a. Under \$16,000
  - b. \$16,000 - \$34,999
  - c. Over \$35,000
  - d. Missing
5. Presence of wood stove or fireplace insert
  - a. Present
  - b. Not present

The criteria were applied in an hierarchical manner, as illustrated in Figure 2. After collapsing cells containing few or no cases, the final stratification matrix contained 128 cells. The sample was then allocated to the cells roughly in proportion to the estimated distribution of the PNWRES-83 participants. Within each cell, the appropriate number of sites was selected using the SPSS-x sample algorithm (SPSS Inc., 1983, pgs. 151-155). This resulted in a "Primary Sample" of 400 dwellings.

Due to the level of commitment required of the participants, a relatively high nonresponse rate was anticipated. Due to the geographic dispersion of the sample, from a logistics and budgetary perspective, it seemed most reasonable to select a specific set of replacement sites. To contain costs, for each Primary Sample Site, the most geographically proximate site satisfying the same stratification criteria was chosen as the replacement.

Due to the limited number of eligible sites in some strata, certain sites served as replacements for two or more Primary Sample sites. As a result, the number of first order replacement sites was somewhat less than 400. Due to the high refusal rates in certain cells, and the need to reject certain sites for potential fire hazard problems, additional replacement sites were selected during the recruiting process. In the end, 811 sites were selected from the 1445 eligible PNWRES-83 sites.

#### PARTICIPATION RATES

Site recruiting was considered complete when the homeowner (and, in the case of rental units, the current tenants) signed an Access Agreement permitting the installation of the data logging equipment (Parker, et al., 1985). As depicted in Table 1, the participation rate for Base Study dwellings in the Primary Sample is nearly 50 percent: 198 of the 400 units.

Since some of the units were never contacted by recruiters, the data exaggerate the nonresponse rate among "Replacement" dwellings. Nevertheless, the nonresponse rates are in general high, which raises questions as to the potential bias of the final sample (Kish, 1965 pg. 535ff).

#### CHARACTERISTICS OF THE FINAL SAMPLE

The characterizations reported here are based on the data from the PNWRES-83. Although more recent data are available for the ELCAP participants, the PNWRES-83 data were used to select the sample. Furthermore, they are the only consistent set of data for the entire sample frame.

To facilitate interpretation, we report the weighted estimates from the PNWRES-83 for the entire PNW region (L. Harris & Associates, 1984c), as well as the unweighted results for the full PNWRES-83 sample. We then report the distributions for the pool of Base Study eligible sites, for the Primary Sample, and, finally, for the Final Sample of participating residences. In some cases, we have also reported the associated standard deviations together with the appropriate measures of central tendency. However, these values are the standard products of SAS (SAS Institute, 1985), and do not take into account the substantial design effects associated with the sampling methodology employed in the PNWRES-83. Hence, the values are illustrative only, and should not be used in any statistical tests.

Despite the high nonresponse rate, the intensive stratification strategy appears to have succeeded in ensuring that the sample is relatively well distributed over the criteria considered critical, such as climate (geographic area), dwelling vintage, and income (Tables 2 through 6). Any substantial differences between the final sample and the full weighted or unweighted regional population appear to be due at least as much to the sample allocation algorithms as to any self-selection on the part of the potential respondents. Thus, the sample departs severely from the weighted distribution over utility ownership type, but conforms rather closely to the distributions of both the "eligible" and the "Primary Sample" sites (Table 3). Similarly, the sampling procedures sought an approximately equal distribution between units with and those without wood heating equipment (Table 6).

Nevertheless, there does appear to be some evidence of systematic bias, even among the basic stratification characteristics. For example, there are fewer older homes and fewer low income households than there are even in the pool of eligible sites (Tables 4 and 5). The participating households also tend to be larger (Table 8), the householders tend to be more well-educated (Table 7), and they tend to own a slightly larger number of appliances (Table 9). While these differences may have been anticipated and reasonable explanations are almost immediately apparent, nevertheless, they suggest a possible bias in the data. The suggestion appears to be reinforced by the comparatively higher ratio of the minimum kWh to maximum kWh per billing period and the higher reported rates of recent conservation actions (Tables 10 and 11). In other words, it appears that ELCAP participants may have a larger "base load" and comparatively smaller "space heating load" than is true of the general population or for the frame of eligible units. Clearly, these sorts of differences could have significant consequences for estimates of end-use shares, estimates of potential conservation, as well as estimates of typical residential load shapes.

#### CONCLUSIONS AND IMPLICATIONS

The ELCAP residential Base Study was originally intended to support Bonneville's conservation assessment and load forecasting responsibilities. In both cases, the full and effective performance of the responsibility involves the derivation of estimates of population parameters. From the outset, however, the design of the residential Base Study acknowledged the inability of deriving such estimates using purely statistical procedures. The frame for the sample was restricted to a subset of the population and the technique used to select replacement sites raises some doubts concerning the randomness of the sample. Finally, as we had anticipated, the low response rate raises questions concerning potential bias; questions which the analysis presented here appears to answer in the affirmative. Although it is perhaps possible to employ post-stratification techniques to derive appropriate case weights, this would only permit the analyst to derive statistically valid estimates for about 25 percent of the population.

As an alternative, the strategy employed in the selection of the Base Study sample was to focus on certain critical criteria and attempt to ensure representation across the full range of each criterion. The data could then be used in an hypothesis testing framework to explain observed variations in the primary dependent variables, i.e., annual, monthly, weekly, hourly electricity consumption.

Based on the brief analysis presented here, it appears that the strategy was, for the most part, successful. There are some weaknesses at the extremes of certain characteristics, including those explicitly involved in the stratification (i.e., housing vintage and income). In retrospect, these categories should have been oversampled in order to ensure adequate representation among the final participants. In the meantime, analysts will want to pay special attention to the effects of these extreme values. Analysts will also want to pay special attention to those several characteristics for which there is some evidence of potential bias.

#### REFERENCES

- Louis Harris and Associates, 1983, PNWRES-83: Sample Selection Activities Report, Portland, OR, Bonneville Power Administration
- Louis Harris and Associates, 1984a, PNWRES-83: Fieldwork and Data Reduction Activities Report and Appendices, Portland, OR, Bonneville Power Administration
- Louis Harris and Associates, 1984b, PNWRES-83: Response Analysis Report, Portland, OR, Bonneville Power Administration
- Louis Harris and Associates, 1984c, PNWRES-83: Construction of Multicomponent Case Weights, Portland, OR, Bonneville Power Administration
- Kish, Leslie, 1965, Survey Sampling, New York, John Wiley & Sons
- Parker, G., E. Pearson and W. Sandusky, ELCAP: The Residential Pilot Study, Portland, OR, Bonneville Power Administration
- SAS Institute, Inc., 1985, SAS User's Guide: Basics, Version 5 Edition, Cary, NC, SAS Institute, Inc.
- SPSS Inc., 1983, SPSS-x User's Guide, New York, McGraw-Hill Book Co.
- Stokes, R. et al., 1985, ELCAP: Papers Presented at the Electric Power Research Institute's Commercial End-Use Metering Workshop, Portland, OR, Bonneville Power Administration
- Tomich, S. and G. Schuster, 1985, ELCAP: Description of the Field Data Acquisition System for Building Energy Monitoring, Portland, OR, Bonneville Power Administration
- Windell, P., forthcoming, ELCAP: Residential Base and Case Studies Sample Design, Portland, OR, Bonneville Power Administration

FIGURE 1

SAMPLE FRAME DEFINITIONS

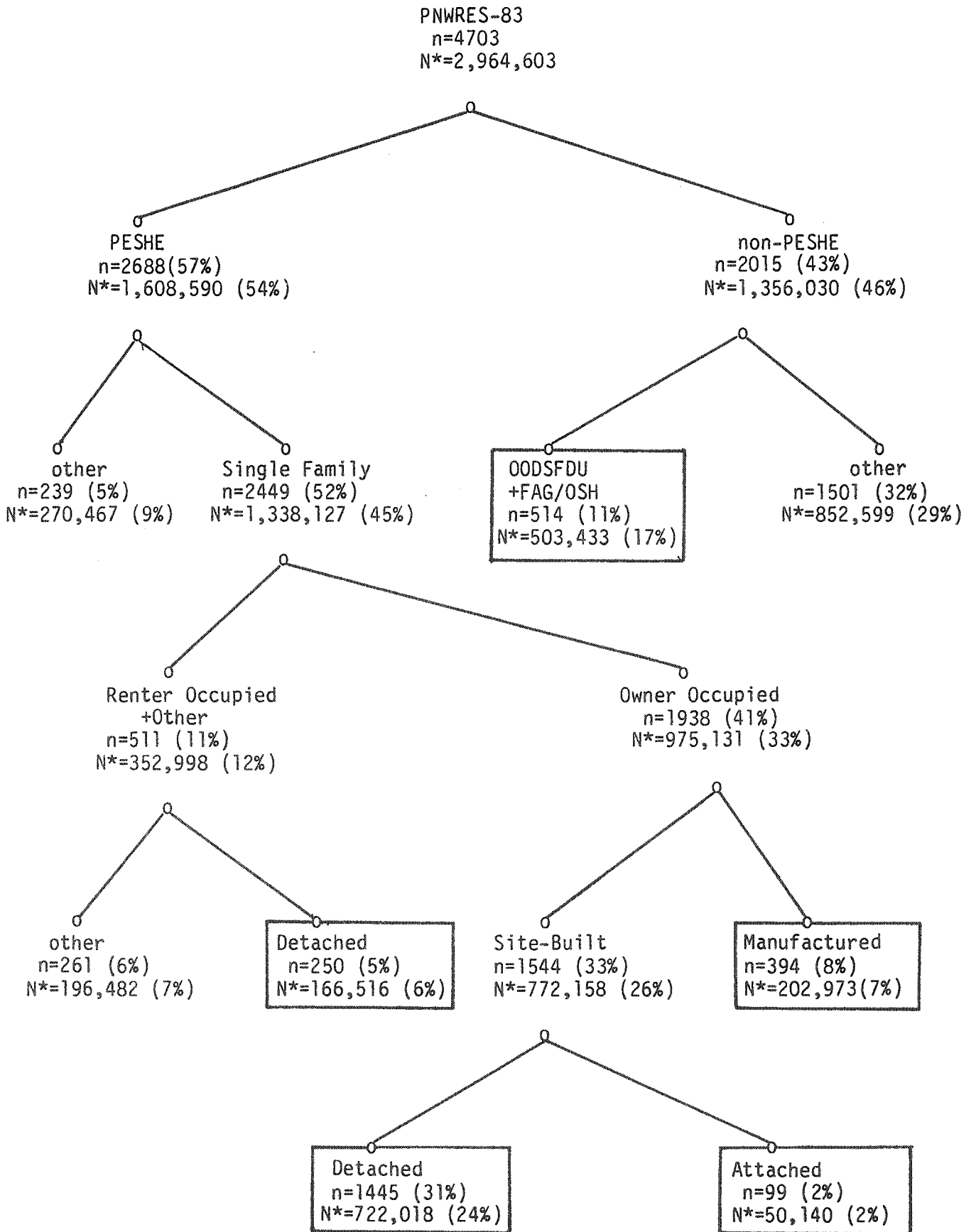


FIGURE 2  
Stratification of Base Study Frame

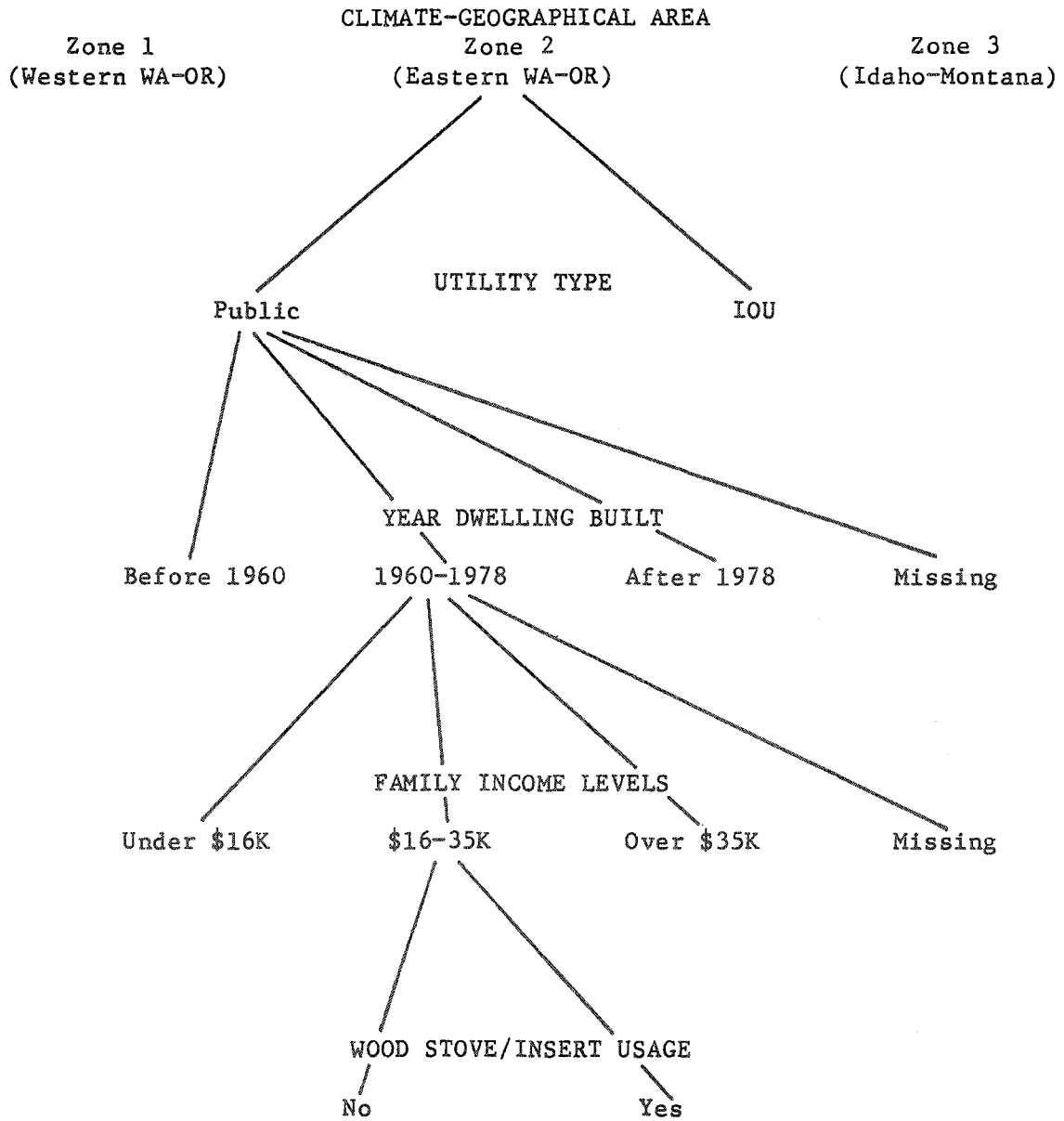


TABLE 1: SAMPLE SELECTION STATUS BY PARTICIPATION

=====

SELECTION STATUS	FINAL PARTICIPATION		TOTAL
	NOT PARTICIPANT	PARTICIPANT	
PRIMARY	202	198	400
REPLACEMENT	337	74	411
	---	---	---
TOTAL	539	272	811

TABLE 2: DISTRIBUTION BY GEOGRAPHIC AREA

=====

	WWA	WOR	EWA	EOR	ID	MT	N
	---	---	---	---	---	---	---
PNW REGION (EST)	39.8	28.2	12.5	4.5	11.6	3.4	2,964.3
PNWRES SAMPLE	18.6	15.9	17.8	15.7	16.5	15.6	4,703
BASE STUDY:							
ELIGIBLE	19.2	14.3	21.6	14.8	15.6	14.5	1,445
PRIMARY SAMPLE	22.0	18.5	20.3	14.3	12.0	13.0	400
FINAL SAMPLE	22.4	18.7	21.3	13.8	13.4	10.4	268

TABLE 3: DISTRIBUTION BY UTILITY OWNERSHIP TYPE

=====

	POU	IOU	N
	---	---	---
PNW REGION (EST)	39.9	60.1	2,964.3
PNWRES SAMPLE	57.8	42.2	4,703
BASE STUDY:			
ELIGIBLE	70.3	29.7	1,445
PRIMARY SAMPLE	70.0	30.0	400
FINAL SAMPLE	69.4	30.6	268

TABLE 4: DISTRIBUTION BY YEAR DWELLING BUILT

=====

	<1950	50-59	60-69	70-77	78-83	MISSING	N
	---	---	---	---	---	---	---
PNW REGION (EST)	29.6	12.2	15.3	25.0	9.8	8.1	2,964.3
PNWRES SAMPLE	31.3	11.3	13.5	27.9	10.2	5.8	4,703
BASE STUDY:							
ELIGIBLE	28.7	12.0	12.0	32.9	12.2	2.1	1,445
PRIMARY SAMPLE	27.3	13.0	14.3	32.0	12.0	1.5	400
FINAL SAMPLE	23.9	11.2	14.6	36.2	12.3	1.9	268

TABLE 5: DISTRIBUTION BY 1982 HOUSEHOLD INCOME

	LT \$8K	\$8K- \$16K	\$16K- \$25K	\$25K- \$35K	\$35K- \$60K	GT \$60K	MISSING	N
PNW REGION (EST)	13.7	22.8	20.1	16.0	11.9	2.6	12.8	2,964.3
PNWRES SAMPLE	14.7	23.2	20.7	15.5	10.4	2.7	12.7	4,703
BASE STUDY:								
ELIGIBLE	7.9	18.1	22.1	19.0	15.0	4.5	13.4	1,445
PRIMARY SAMPLE	8.3	17.5	21.8	20.0	15.8	5.0	11.8	400
FINAL SAMPLE	5.6	14.2	25.0	20.5	20.1	4.5	10.1	268

TABLE 6: DISTRIBUTION BY PRESENCE OF WOOD STOVE

	PRESENT	NOT PRESENT	N
PNW REGION (EST)	29.3	70.7	2,964.3
PNWRES SAMPLE	40.3	59.7	4,703
BASE STUDY:			
ELIGIBLE	46.4	53.6	1,445
PRIMARY SAMPLE	44.0	56.0	400
FINAL SAMPLE	45.9	54.1	268

TABLE 7: DISTRIBUTION BY HOUSEHOLD EDUCATION

	LT H S	H S GRAD	TRADE -VOC	SOME COLL	COLL GRAD	POST GRAD	MISS	N
PNW REGION (EST)	12.3	24.6	7.6	25.8	14.3	14.5	0.9	2,964.3
PNWRES SAMPLE	13.7	26.3	7.4	26.2	13.1	12.5	0.7	4,703
BASE STUDY:								
ELIGIBLE	9.6	22.6	8.4	27.1	14.4	17.2	0.6	1,445
PRIMARY SAMPLE	9.5	19.3	7.8	26.0	15.5	21.3	0.8	400
FINAL SAMPLE	6.7	17.9	8.2	26.9	18.7	21.3	0.4	268

TABLE 8: HOUSEHOLD AND DWELLING UNIT SIZE

	NUMBER OF RESIDENTS			SIZE OF DWELLING (SQ.FT.)	NUMBER ROOMS	N
	TOTAL	LT 6YO	GT 64YO			
PNW REGION (EST)	2.68	0.24	0.33	1461.0	6.23	2,964.3
PNWRES SAMPLE	2.75	0.27	0.35	1485.2	6.30	4,703
BASE STUDY:						
ELIGIBLE	2.91	0.26	0.37	1796.8	7.12	1,445
PRIMARY SAMPLE	2.98	0.27	0.37	1824.2	7.18	400
FINAL SAMPLE	3.15	0.33	0.29	1890.0	7.25	268
STANDARD DEVIATION	(1.51)	(0.62)	(0.64)	(801.3)	(2.08)	

TABLE 9: HOUSEHOLD APPLIANCE INDEXES

=====

	BASIC	SMALL	LARGE	TOTAL	N
	-----	-----	-----	-----	-----
PNW REGION (EST)	6.48	2.03	0.45	8.97	2,964.3
PNWRES SAMPLE	6.85	1.90	0.56	9.31	4,703
BASE STUDY:					
ELIGIBLE	7.71	2.11	0.72	10.54	1,445
PRIMARY SAMPLE	7.70	2.07	0.70	10.47	400
FINAL SAMPLE	7.97	2.23	0.80	11.01	268
STANDARD DEVIATION	(1.94)	(1.25)	(0.89)	(3.04)	

TABLE 10: CONSERVATION ACTIONS TAKEN 5/80 - 5/83

=====

	INSUL. ROOF/ CEILING	STORM WINDOW	WINDOW PLASTIC	WATER HEATER WRAP	LOWER WATER TEMP.	LOWER SPACE TEMP.	N
	-----	-----	-----	-----	-----	-----	-----
PNW REGION (EST)	21.4	19.8	17.3	32.2	37.5	32.5	2,964.3
PNWRES SAMPLE	22.5	20.0	20.7	32.3	37.0	29.2	4,703
BASE STUDY:							
ELIGIBLE	29.6	25.5	16.2	38.4	41.7	32.8	1,445
PRIMARY SAMPLE	31.3	26.5	14.3	38.5	44.8	34.8	400
FINAL SAMPLE	31.7	28.4	14.6	43.7	45.9	35.1	268

TABLE 11: 1982 ELECTRICITY CONSUMPTION

=====

	TOTAL kWh	kWh/ SQFT	kWh/ DAY	kWh/ CAPITA	minkWh/ maxkWh	N
	-----	-----	-----	-----	-----	-----
PNW REGION (EST)	13,125.4	12.1	55.2	6,841.0	.410	2,964.3
PNWRES SAMPLE	14,446.8	12.7	60.5	7,178.6	.360	4,703
BASE STUDY:						
ELIGIBLE	20,746.1	14.2	68.5	9,369.1	.286	1,445
PRIMARY SAMPLE	21,958.9	14.3	72.1	9,498.3	.296	400
FINAL SAMPLE	22,834.2	14.3	74.2	8,837.6	.304	268
STANDARD DEVIATION	(11,580.9)	(9.2)	(141.8)	(5,528.8)	(.201)	