General Principles for Determining
Sample Sizes for Appliance-Saturation Surveys

1. Sample Size

While each utility must determine, itself, appropriate sample sizes for appliance-saturation surveys, the following are some general principles for determining sample sizes.

The heating, water heating and cooling saturation rates are critical components in obtaining population estimates of residential load profiles. In building up to aggregate LDC residential load profiles, aggregate electric heating, electric water heating, air conditioning and base load profiles are first assembled. To obtain these four profiles, basically an estimate is needed for the last column of the following table.

Electric Heating	Presence of Electric Water Heating	Air Conditioning	LDC's Proportion
Yes	Yes	Yes	
Yes	Yes	No	
No	Yes	Yes	
No	Yes	No	
No	No	Yes	
No	No	No	

For some utilities the population estimates of some of the components of the last column could be as small as 5%. If the population saturation rate is that low, it is recommended that we should be within $\pm 3\%$ of the true saturation rate 95% of the time. The $\pm 3\%$ is equivalent to a standard error of 1.53%.[1] This has been taken into account in Table 1.

Where a LDC has a seasonal or a density subclass, a separate appliance-saturation survey is recommended for each subclass. In the following discussion, for a LDC with a subclass, the population size along with suggested sample sizes correspond to that of the subclass.

The following table considers two types of sampling, random sampling and proportionate sampling. In calculating the proportionate sampling numbers it has been assumed that proportionate sampling should reduce the variance by a factor of about 0.75. After providing the sample sizes for residential classes of various sizes using the two methods, brief instructions on selection of samples will be provided. This is particularly relevant for proportionate sampling.[2]

[1] Standard errors in the range of 5% are not meaningful if we are constructing a 95% confidence interval of ± 1.96 standard errors.

[2] The calculations are based on formulae provided by Kish (1966; Survey Sampling, John Wiley & Sons Inc.)

Table 1

Population Size	Sample Size for Random Sampling	Sample Size for Proportionate Sampling
250,000	1063	798
200,000	1062	798
100,000	1057	795
50,000	1046	788
40,000	1040	785
20,000	1014	770
10,000	965	742
5,000	880	690
2,500	748	607
600	384	343

For random sampling, customer records would be randomly drawn using a random number generator and appropriate calls would be made. Allowing for non-response by potential participants or not reaching people when called, an initial list may have to be five times as large as the required sample.

Proportionate sampling involves fewer samples. The reason is that this method takes advantage of possible homogeneity within strata. For example, one might presume that the majority of households with greater than 40,000 kWh are electrically heated. For proportionate sampling, based on annual billing kWh, one would construct a frequency distribution of percentage of the population in each of the strata. The following are examples of strata used in previous load research.

kWh range	Population Proportion
0 – 6,000	n_1
6,000 – 10,500	n_2
10,500 – 16,500	n_3
16,500 – 24,000	n_4
24,000 – 40,000	n_5
> 40,000	n_6
	N

So, if from above, the total sample size is 795 (for a population of 100,000), a random sample of $\frac{n_1}{N}795$ from stratum 1 would be sampled, a random sample of $\frac{n_2}{N}795$ would be sampled from stratum 2, and so on. Again, because of possible non-response in each stratum, an initial sample from each stratum would have to be at least five times the required sample.

Where a LDC, with best effort, undertakes the above but does not meet the target sample sizes, it will mean that sample estimates will have a larger dispersion than being within $\pm 3\%$ of the true saturation rate 95% of the time. However, this estimate is better than no estimate. This "best effort estimate", along with the kWh distribution of survey participants, can be combined with billing data of the class to improve the reliability of the estimate. Therefore, some survey data from a sample where the target sample size has not been met, is still extremely valuable.

To get the participation rate up for the survey, incentives such as prizes or gift certificates can be very helpful.

Depending on the method used, there may be concern about possible biases of the resulting sample. After administering the survey, a check on its representativeness relative to say an annual billing kWh distribution of the class can be performed. If there are biases, a correction can be done that reweights part of the sample based on the population kWh distribution of the class. This can be done without any additional collection of survey data.

2. Determining whether Two Subclasses Require Separate Saturation Surveys

Suppose we want to test whether two subclasses within a utility have the same saturation rates. Identical saturation rates should lead to identical load profiles for the subclasses, given that the end-use profiles would be identical.

Given that we may not have any prior sample information about saturation rates, there is not a foolproof way of testing for identical saturation rates. A rough check would be to look at the equality of corresponding annual kWh distributions. Here we would be using the distribution of annual kWh as a proxy for identifying various saturation rates. But, as can be appreciated, there is not a one-to-one correspondence between saturation rates and annual kWh distributions.[3]

Given the above caveats, here is a suggested way of testing whether two distributions of annual kWh are equal. Let one distribution (j=1) correspond to annual billing kWh of the first subclass. Let the other distribution (j=2) correspond to annual kWh of the second subclass.

Divide the billing records into six categories, $i = 1,2,...,6$. Record the number of customers (X_{ij}) in the ith category. Include only respondents with a complete year of kWh billing data. In other words, assemble the following table.

[3] Comparing monthly kWh distributions would be more precise; nevertheless, due to billing cycle issues and due to billing kWh not always corresponding to actual usage of kWh, a true monthly comparison may not be possible.

Table 2

kWh range	Subclass, j=1	Subclass, j=2
0 – 6,000	X_{11}	X_{12}
6,000 – 10,500	X_{21}	X_{22}
10,500 – 16,500	X_{31}	X_{32}
16,500 – 24,000	X_{41}	X_{42}
24,000 – 40,000	X_{51}	X_{52}
> 40,000	X_{61}	X_{62}
Total	n_1	n_2

The null hypothesis to be tested is

$$H_0 : p_{i1} = p_{i2} \ \ for\ i = 1,2,...,6$$

where p_{ij} correspond to the proportion of population j in category i.

For the implementation of the following test we will assume that we are drawing a sample of N_1 from each population 1 and N_2 from population 2. N_1 and N_2 are determined from column 2 (Sample Size for Random Sampling) of Table 1. To test the above hypothesis, the following χ^2 test statistic can be used.

$$\chi^2(k-1) = \sum_{j=1}^{2} \sum_{i=1}^{k} \frac{\left\{ X_{ij}^{n} - N_j \left[\frac{X_{i1}^{n} + X_{i2}^{n}}{N_1 + N_2} \right] \right\}^2}{N_j \left[\frac{X_{i1}^{n} + X_{i2}^{n}}{N_1 + N_2} \right]} \text{ where } X_{ij}^{n} = \frac{N_j \bullet X_{ij}}{n_j}$$

and, in the example above, with 6 categories, k=6.

Now compare $\chi^2(5)$ with 11.1 (5% critical value). If $\chi^2(5)$ is greater than 11.1, then H_0 is rejected. Thus individual sets of saturation rates must be collected for each class.

If the null hypothesis H_0 is not rejected, then we can presume the two subclasses would have the same load profiles and the same saturation rates. This would imply that one subclass could borrow saturation rates from the other subclass or that the two subclasses could be pooled and sampling from the pooled population (with sample sizes determined appropriately from the above table regarding sample sizes for random and proportionate sampling) forms the basis of estimated saturation rates.

3. Sharing Saturation Rates among Similar Utilities

An extension of the above could be used to test whether a residential class in one utility resembles the residential class of a neighbouring utility (e.g., a utility in the same geographic region, or utilities in different regions but in very close proximity to each other) with a class of similar or larger customer size.[4] Here caution must be exerted. The objective still remains to get the appropriate saturation rates for the utility and class under consideration. However, recall that the annual kWh distribution is being used as a proxy for saturation rates. As geographic distances start increasing, kWh strata boundaries mapping to annual usage of electric heating or air conditioning, start changing. Even annual usage of electric water heating can start varying as significant differences in geography are noted.[5] Thus, although the above tests are based on predetermined kWh strata, these strata start loosing meaning with respect to correspondence to saturation rates, as geographic distances between utilities grows. In other words, the annual kWh proxy for saturation rates starts losing its applicability.

Nevertheless, if this extension across utilities is pursued, just replace the two "subclasses" with "two classes of neighbouring utilities". And, here, if the null hypothesis is not rejected, then a residential class in one utility could borrow saturation rates from a neighbouring utility of similar or larger size. Another possibility is for the neighbouring utilities, if they satisfy the above criteria, jointly conducting an appliance saturation survey with the population now being the pooled populations of the neigbouring utilities. If the latter pooling is done, the neighbouring utilities should be of roughly similar customer size.

Dean C. Mountain
March 31, 2006

---

[4] For example, limit the borrowing of saturation rates to utilities with suggested sample sizes (see Table 1) within 97.5% of the target sample size of the utility under consideration. For example, a utility with a residential class of size 200,000 (with a suggested sample size of 1062) could borrow saturation rates from another utility with a residential class of 40,000 (with a suggested sample size of 1040) or above, provided the other conditions (e.g., neighbouring utility criterion, and χ^2 test criterion) are satisfied.

[5] For example, some utilities are drawing their water from the lake as opposed to other utilities drawing their water from wells. Just the temperature difference from these two water sources can translate into different kWh usage patterns for electric water heating.