Mercury Thermostats: Estimating Inventory and Flow from Existing Buildings

Technical Approach Summary

Project Approach prepared for:
Mark Tibbetts, Thermostat Recycling Corporation

Prepared by:
Lisa A. Skumatz
Skumatz Economic Research Associates, Inc.
762 Eldorado Drive, Superior, CO 80027

March 1, 2009

1. Understanding of the Project

A number of states are passing mandatory mercury thermostat collection laws and associated collection goals. However, the goals are imperfectly set, as information on the number of mercury thermostats in either commercial or residential buildings is largely unknown. Developing targets, planning for handling and diversion/ disposal, and investing in the design of programs are all improved with more reliable information on the inventory of mercury thermostat, their removal rates and practices, and the actors involved.

With the exception of one statistical study (King County, commercial sector) most of the goals were set using "rule s of thumb" or very crude assumptions. For example, the State of Iowa asserts an assumption that each home has 1.25 thermostats for its goal-setting computations.

Our review of the needs for this project and potential analytical techniques makes it clear that the advanced techniques used for inventorying and measuring lifetimes for *energy efficiency* equipment are directly applicable to this project, and provide strong statistical methods to develop appropriate estimates. The methods involve large-scale survey data collection[1] and statistical / regression analysis methods[2] designed for just this type of application. We have considered a variety of methods, and we propose a study that adapts and streamlines some of the King County methods to provide reliable estimates of the following information:

- Number of mercury thermostats[3] in place in residential buildings and separately in commercial buildings - useful for an inventory and to predict the lifetime flows of equipment;
- Number of annual removals of mercury thermostats from residential and commercial buildings (separately) – useful for setting annual goals;
- Practices in removing thermostats (and destinations), actors involved in decision-making and removal, and factors influencing removal rates in commercial vs. residential buildings –.useful for identifying intervention points, possible program / capture designs to most effective (and cost-effectively) redirect the equipment to proper disposal, etc.

Our research design and basic task descriptions follow. We embed lessons learned from our previous work on these types of projects as applied to energy efficiency work throughout the description to add context to the approach.

[1] SERA has conducted scores of Residential appliance saturation surveys, and commercial end use surveys for utilities across the US, inventorying all kinds of energy using equipment – and their usage patterns -- in buildings. We have also conducted surveys on building remodeling / change out and equipment removal and the statistical analysis methods required to predict measure retention / measure lifetimes. This work has been conducted for the Bonneville Power Administration (we developed / pioneered the techniques that are now "standard practice"), for the California Public Utilities Commission (CPUC), for Southern California Edison, Pacific Gas and Electric Company and others. We conducted the review of the measure lifetime figures from California's DEER database update, and are just completing the study for the CPUC on "best practices" for conducting energy measure retention estimates for the State.

[2] SERA has used advanced statistical analysis techniques (hazard functions, decay patterns) to measure retention / measure lifetimes and removal patterns and effective useful lifetimes (EULS) (as mentioned above), SERA has expertise in identifying missing data bias in surveys, and helped develop "hot deck" imputation methods to input missing data to reduce bias in energy-equipment related surveys (see peer-reviewed article in The Energy Journal). This provides a strong foundation for estimating the equipment and its removal patterns.

[3] Information on in-duct "sensors" is more complicated to obtain, likely requiring an on-site survey. Given the expense, and the relatively small amount of mercury contained, data on sensors is not included in this effort.

2. Approach and Task Descriptions

Task 1: Preparing the Survey

SERA has conducted more than 100 large-scale surveys for both residential and commercial energy-related equipment inventories and similar studies. Over time, techniques have evolved, and while we originally used paper surveys, we evolved through phone surveys, and most recently, we have found that web surveys are the most effective and cost-effective. [4]

Postcard outreach to web survey: Our plan for both the residential and commercial sectors is to send postcard notifications providing a link to a web survey. Filling out the survey enters the respondent into a lottery to win one of several MP3 players[5] or similar. We will purchase a neutral website upon which to post the survey (e.g. thermostatinfo.org etc.), or use SERA's website (as we have done for many previous surveys). Users will go to the website and click on the labeled button to launch the survey. A web survey provides several advantages.

- Inexpensive and very speedy data collection
- Respondents can fill out the survey night or day – at their convenience – which turns out to be especially convenient for businesses;
- Skip patterns are automatic - so errors do not arise and the survey can be shorter or longer based on specific responses;
- Drawings / pictures can be incorporated for clarification of issues;
- Data is automatically entered into the computer – no separate keypunch entry costs (and errors).
- On-going data checking – we can look at responses after the first few and make adjustments to the survey to correct for anything that seems unclear or to probe on issues; and the data can be analyzed in "real time" and the survey "left active" and analyzed again when more responses have been received.

To allow flexibility, we will also provide a toll-free number on the postcard and administer the survey via phone to anyone preferring that option (usually elderly homeowners, but also some businesses). To increase response and reduce bias, we propose to send an initial postcard, and then several weeks later, a reminder postcard.[6]

Sample source and sampling design: We will use random, and stratified random samples for the survey, using lists purchased from third party vendors. The approaches for the residential and commercial sectors are presented below.

- Residential: We propose to purchase a random sample of home addresses from one of several sources (we often use InfoUSA). We will use skip patterns embedded in the survey so questions appropriate for single family residential vs. multi-family residential vs. commercial respondents can answer appropriate questions. We do not expect to stratify by age of home, but will expect to stratify by geographic area to assure we gain large enough responses from lower-density areas that may have different development (and thermostat replacement patterns) than urban areas. Regarding age of home, given the period over which these thermostats were in place, replacement / remodeling pattern uncertainties, and our interest in developing forecasting models, it is probably best to include all years.

---

[4] Response rates on the order of 15-25% and higher have been received from our previous postcard/web surveys even when the return address is an agency with which the respondent has no relationship (that is, often response rates can be higher when received from your "own" utility). This survey will not have that relationship.

[5] We have found this prize works amazingly well for both the residential and commercial sector. The cost of even several dozen MP3 players is still many times less than the cost of a large scale phone survey (other options include gift certificates, checks, and other options). We have gotten strong response rates from this approach, and the bias from web access may not be dramatically different from problems introduced in phone surveys from unlisted numbers and cell phone only households.

[6] An adaptation of the "Dillman" method for reducing survey bias.

- **Commercial**: The proper stratification method for the commercial sector is almost always by size and business or building type.[7] However, acquiring that data ahead of time when your client is not the utility that has such data (from energy bills and other records) is complicated. We can, however, easily purchase data for a sample of businesses by employment size from third-party vendors. Distributions of businesses (or buildings) by size are far from even – there are many more small firms / buildings than large ones. However, energy use – and we hypothesize possibly the presence of thermostats[8] – may also be disproportionate. If there is any reason to believe that there are differences in the responses from large vs. small buildings (in thermostats, in response rates, in decision-making or other topics of our analysis, or if we want to be able to present information separately by size or building) we will want to assure we have reasonable response numbers from these large buildings. Thus, we will want to send out a larger proportion (but not necessarily larger total number) of postcards to this group. This stratification technique can improve the quality and robustness of the database. Because small buildings or businesses tend to be more homogeneous, a smaller percent (although still large number) of respondents can be included in the sample. Generally, the percent of sample out of the census listing increases as you increase in size, providing explanatory power for the types of variables of interest in this study.

Questionnaire content: Data are collected that 1) provide key analytical data of interest, and 2) provide demographics or "firm-o-graphics" that allow us to identify whether there is potentially bias in responses, and thus the potential for bias in the data (bias detection and correction is addressed in the next task). This supporting data must be asked in a way that allows us to match response proportions against the initial sampling frame database or against published "census" or other data. We plan to include the following question topics in our survey.

- Whether there are any thermostats in building
- Type of thermostat (work with TRC on how to identify types – especially those with mercury)
- How many thermostats in building of the type(s) of interest; whether each is still functioning (&)[9]
- Year built (*); year equipment installed (&) (same for OEM)
- Year remodeled (*); year remodeled that removed / involved thermostat (&); reason for removal for any thermostats removed (&); type of remodel (soft remodel / tenant improvement vs. larger effort); Expected timing of next remodel / next remodel involving thermostat
- Square feet(*)
- Predominate use of the building(*)
- Who removes thermostat; destination of removed thermostats, actors involved in decision-making and removal
- Factors influencing remodeling and thermostat removal rates (in commercial buildings especially) like economy, energy programs, requirement payback for equipment change out
- Other information useful for identifying intervention points, possible program / capture designs to most effectively (and cost-effectively) redirect the equipment to proper disposal, etc.
- Demographic data: type of home (SF detached, attached, MF, etc.); number of residents, square footage, number of bed and bath rooms (predictor for square footage); education of head of household and/or income; zip code[10], other
- Firmographic data: type of business, main activity at the site; single firm or chain / central decision-making, square footage, number of employees, other.

---

[7] The remainder of this discussion assumes that the survey entity is a business, not a building. If we can identify a source for "building" addresses with an associated respondent firm, we can substitute that if desired.
[8] Although the information from the King County study indicates otherwise – they did not find significantly more thermostats in large buildings; however, it is important to identify whether this finding is corroborated.
[9] The group of 4 features marked with asterisks (*) were identified as strong explanatory factors for the King County work. The group for data collection elements marked with ampersands (&) represent those needed to support measure retention / removal analyses.
[10] Zip codes will allow us to check for geographic bias in the results and weight to improve representation if needed.

Sample Size: It is natural to assume that that sample size is directly related to "coverage" of the survey – that is, you would think you would need many more surveys to represent a state as opposed to representing a city. Actually, that is not true, and is illustrated in the table below. The number of responses needed to represent a city of 10,000 and a state of 10,000,000 at +/-5% accuracy with 95% confidence[11] is not very different. No matter how big the total households (within a fairly large band), the sample sizes needed are similar.[12] Regarding the question of how many surveys it takes to achieve higher levels of accuracy, statistical sample sizes are far from proportional; double the accuracy requires more than twice the sample. The budget implications for sample size options is provided in the budget section of this proposal. Generally, however, more data (rather than less) help predictive modeling work, making estimates more robust.

Table 1: Computation of Sample Sizes and resulting accuracy / confidence in responses

IF your population (homes or comm'l bldgs) is…	95% confidence		90% confidence	
Computed responses needed for accuracy of…	+/-5%	+/-10%[13]	+/-5%	+/-10%
100	79	49	73	40
1,000	278	88	213	63
10,000	370	95	263	67
100,000	383	96	270	68
1,000,000	384	96	271	68
10,000,000	384	96	271	68

We will prepare a draft residential postcard, draft commercial postcard, and draft questionnaire for TRC review. We will then incorporate comments and finalize the documents. We will conduct test runs on the survey instrument to assure proper functionality. We will prepare a detailed sampling plan, and purchase the data necessary to support the mailing. We will have the postcards printed and mailed. We envision black and white postcards printed on colored cardstock.[14] We will monitor responses and keep TRC posted on a weekly basis on the number of responses obtained. We will prepare the second mailing for about two weeks after the first mailing. A summary of basic responses can be provided weekly with the "count".

Task 2: Bias Reduction and Analysis

Bias Reduction: Before we conduct any analyses, we will review the data for evidence of bias. There are two types of bias, each addressed below.

- Response bias, or systematic non-random patterns in who responded to the survey. For example, our sampling plan may have been designed to achieve 20% of our responses from small buildings, but we find we are only getting 10% from them, or patterns of data disproportionately for newer vs. older buildings (compared to the purchased sample or to data from national or regional studies). Alternatively, we may compare our "age" distribution for household responses and find we have too many from the "senior" age groups, or check county or zip and identify geographic response bias. These elements can (and will, if found) be corrected by revising the "weights" assigned to each survey. Simply put (and simplified!) if we got perfect responses from all

[11] The statement of confidence level is standard practice; however, it is simplified and slightly misleading. It states the accuracy with which you would predict a 50% response (e.g. male / female, yes/no) given random responses totaling the number given from the sample, compared to the answer from the population. It does not predict the accuracy of an answer of, say, "9" from among a number of categories, etc. However, it serves as a benchmark for higher vs. lower accuracy sample sizes.

[12] This is why Nielsen television ratings or Gallup election polls can use nationwide samples of 1200 and get accuracy nationwide.

[13] Also equals +/- 7% at 90% confidence.

[14] Deviations from this will have small budget implications. However, the response rates we noted are from the proposed black and white option.

categories, but twice as many from the elderly, their responses might be weighted as ½ each and then included in the database.

- Patterns in Missing data: The other source of error is in missing data within a survey. For example, younger households may not know when the home was remodeled for a thermostat replacement, and a higher percent leave this answer blank. If this is systematic, it can lead to problems estimating the models. We use "hot deck"[15] imputation methods to "impute" or put in replacement data – but drawing from our sample in doing so. We find responses from "similar" households (same age group, region, house type, etc.) and randomly assign or fill in the response from a "similar" household. It is similar to "predicting" the response and filling that value in, but it has better statistical properties.

We will conduct two key types of analysis:

Predicting "Count" of Thermostats in place: We will use the data from the survey[16] and conduct an analysis similar to that from King County, to predict thermostats in place. We will develop predictive regression models[17] fitted separately for the residential vs. commercial sector to estimate the total number of mercury thermostats in place in the residential sector and in the commercial sector. This will require data we gather throughout the survey on number of thermostats present in the building and the array of associated "explanatory" data for the buildings. These results are then "applied" to the population to predict total number of mercury-containing thermostats in the population. Error and confidence bands can be associated deriving from the sampling as well as the "fit" statistics associated with the models (R-squared, F statistics, t-statistics, etc.).

Predict lifetime and annual "flow" of thermostats out of buildings: We plan to use methods similar to the California Public Utilities Commission (CPUC) "Protocols" for estimating Effective Useful Lifetimes (EULs)[18] for energy efficiency equipment as a key element in predicting the proportion of mercury thermostats in place that will flow out of buildings each year into the future. This analysis is used to predict measure lifetimes and removals of all kinds of energy using equipment, estimates "hazard" functions that relate to the decay and removal of equipment based on a combination of ceasing to function and removal due to remodeling / replacement. It uses the lifetimes of measures that have already been removed to predict how long the remaining equipment will last, and provide an annual flow of removals in percentage terms. We will multiply times the equipment in place (identified from the King County analysis) to compute the actual number expected to be removed each year from the residential vs. the commercial sectors. Using statistical models, we can swap in different decay functions (log, exponential, gamma, etc.) to identify the best "fit". The models can be compared on the basis of a number of statistical "fit" measures similar to R-squared and F-statistics. Thermostat lifetimes have rarely been examined in the measure lifetimes literature, so this will be a valuable addition, and data that are not available elsewhere. The data are easily collected at the same time as the data required for the "King County-type" analysis described above.

We will provide these models and the predictive results as key outputs of the analysis. In addition, we will analyze the data on "influencing" factors, actors, and removal behaviors and destinations to identify whether there are any potentially fruitful ideas for: how to influence the destination of removed equipment toward preferred management /

---

[15] See article by Ong, Holt, Barnes, and Skumatz in "The Energy Journal", 1989.

[16] To place this work in context, we have conducted similar equipment saturation surveys in the residential or commercial sector for more than a dozen utility projects.

[17] King County's analysis of commercial buildings found year built, remodel year, square feet, and predominate building use to be key independent variables

[18] For details on this approach, see Skumatz, "Best Practices in Measure Retention and Lifetime Studies: Standards for Reliable Measure Retention Methodology Derived from Extensive Review", Proceedings of the IEPEC Evaluation Conference, 2005, Brooklyn NY., and similar studies in EEDAL 2006 (London), ACEEE 2002, 2000, 1994, IEPEC 1995 and elsewhere. A detailed review of more than 100 measure retention studies is found in work by Skumatz for the California Public Utilities Commission (CPUC), "Review of Retention and Persistence Studies for the California Public Utilities Commission", 2004, and updates for SCE including "Revised / Updated EULs based on Retention and Persistence Studies Results", 2005 and 2008.

disposal options; how to educate owners or removers about proper treatment; and other issues; factors affecting removal and disposal; geographic patterns and differences, etc. We can identify some of these policy and programmatic issues with TRC at the start of the project.

Task 3: Reporting

This reporting task is focused on two main activities:

- Periodic project management reporting with TRC, and
- Preparation of the final report

For interim reporting, we will provide a monthly written progress report in association with the invoice, and regularly scheduled phone progress meetings every other week with the project manager. The monthly written reports will include progress, planned activities next period, schedule issues and data or information issues holding progress, problems encountered and possible solutions, and important deviations from budget or schedule.

We will provide a draft report for review and comment, and prepare the final report incorporating comments. The report will include:

- Executive summary emphasizing major findings, implications, and recommendations. This is designed to be readable by lay audiences, and should be separable / stand-alone in nature to allow wide distribution.
- Background or introduction, including discussion of the results of the analyses
- Framework and Methodology, providing a description of the research approach, primary and secondary data collection methods, and description of the method / models / approach used for computing the results, etc. Some detailed information may go to the Appendix
- Summary, Implications, and Recommendations, providing a discussion of the major results / issues / implications / recommendations arising from the research.
- Appendix, documenting the data, sources, survey instruments, references, and other documentation related to the conduct of the project.

The Final Report and the presentation will be submitted in hard copy and electronic format. Electronic copies of all raw and analyzed data will also be provided to TRC (with any identifying data removed).

3. Project Staffing and Qualifications / Experience

The key staff on the project includes:

- Dr. Lisa Skumatz. She is a Ph.D. economist and Principal of SERA. Dr. Skumatz has 30+ years of experience in surveys, energy equipment inventories, forecasting, measure decay, and integrated planning/ evaluation for energy programs, as well as recycling / waste management programs. She will be SERA's project manager and will be responsible for organizing the work, conducting the survey design and statistical modeling work, and writing the memos and reports. She will be assisted by Mr. Freeman.
- Juri Freeman. Mr. Freeman is a project manager at SERA. He has conducted scores of survey and analysis projects, and has developed web surveys for about two dozen projects. He has assisted Dr. Skumatz on scores of projects collecting and analyzing data from surveys, and writing analytical reports. He is also administrative lead and back-up project contact.

Dr. Skumatz, Principal of SERA brings more than 30 years of experience in integrated planning, best practices, and evaluation to this project. She has conducted surveys, measure lifetimes, forecasting, integrated planning, best practices, and evaluation of energy efficiency (EE) programs across all customer sectors for more than 30 years. Clients for these assignments include all the IOU utilities in the State of California, Alliance to Save Energy, CPUC, Wisconsin Focus on Energy, Xcel Energy, BCHydro, NYSERDA, and scores of other clients across the US and internationally. She has conducted surveys and analysis for appliance and commercial equipment saturation surveys (inventories) since 1988. She was responsible for the residential appliance surveys for Pacific Gas and Electric Company for three years. Clients for appliance saturation surveys, survey design, and sampling work include (residential and commercial) include: Puget Sound Energy, Northeast Utilities, Seattle City Light, Bonneville Power Administration, CPUC, PG&E, Washington Water Power, Southern California Edison, Midwest Energy Efficiency Alliance, NYSERDA, Sierra Pacific, and many others. Lisa/'s work in Measure lifetimes dates from the first large-scale measure retention survey and estimation work in 1990. She led the data collection and survey design work and the analysis tasks. She developed the statistical analysis methods that now serve as "standard practice" reflected in the state-wide protocols for this kind of work in California and elsewhere. She has served s the measure life time / retention expert on projects for the State of California (CPUC), and elsewhere. Measure lifetime analysis clients include: California Public Utilities Commission, CIEE, Southern California Edison, Bonneville Power Administration, Northeast Utilities, CCIG, Puget Sound Energy, and others. She has conducted forecasting work for PG&E, CPUC, Bonneville Power Administration, among others. Lisa's Ph.D. is in econometric modeling from The Johns Hopkins University, and has about 75 publications in energy conferences, proceedings, and journals.

- *Project Role: Design of survey questions and sampling, bias reduction analysis, statistical modeling for predicting number of mercury thermostats in place in residential and commercial buildings, and the "flow" of thermostat disposed / removed on an annual basis; walkthrough, draft and final report.*

Juri Freeman, Senior Research Analyst. Mr. Freeman manages the design and basic analysis of SERA's web surveys, and is an experienced and skilled interviewer. He has analyzed the results of about two dozen surveys for energy and city clients, and has written repots for clients in the US and internationally. He has extensive experience in measure lifetime analysis work, and has conducted projects for CPUC and Southern California Edison. He helped review and assess more than 100 measure lifetime studies conducted in the State of California He has a degree in Environmental Economics from Bucknell University.

- *Role: Preparation of postcards, preparation of web survey, administer phone surveys (with back-up staff), updates on survey responses, analysis of survey, draft sections of report, review / edit report.*