Review Comments on Proposed Adoption of Survey Approach and Methodology to Estimate Mercury Thermostats Becoming Waste in California

Submitted by:

Catherine M. Crespi, Ph.D.
Associate Professor in Residence
Department of Biostatistics
Fielding School of Public Health
University of California, Los Angeles

Overview of Review Comments

This report provides my review of the approach and methodology of the study "Mercury Thermostats: Estimating Inventory and Flow from Existing Buildings." The California Department of Toxic Substances Control (DTSC) proposes to use the results of this study to develop "… a methodology for … [calculating]… the number of out-of-service mercury-added thermostats becoming waste annually." This is part of DTSC's development of regulations to implement the Mercury Thermostat Collection Act of 2008.

I have reviewed the materials provided and considered them in light of the three scientific conclusions to be addressed. My overall opinion is that the study approach and methodology were reasonable and appropriate insofar as estimating the inventory and annual flow of mercury-added thermostats from residences in California are concerned. The study had limitations in terms of estimating the inventory and annual flow from non-residential buildings. These limitations include:

1. Lack of consideration of the great diversity of non-residential buildings, especially non-commercial buildings and very large entities with multiple buildings (e.g., airports, universities, hospitals, industrial complexes, prisons), in the sampling plan;
2. Low response rate from businesses;
3. Limited number of validation site visits to businesses;
4. Survey questionnaire not well designed to capture data from businesses with complex facilities and/or multiple buildings.

While these limitations lessen the reliability of the estimates for thermostats from the non-residential sector, this concern is mitigated by several factors. It appears from overall numbers that the inventory and flow of mercury-added thermostats out of buildings is dominated by residences, which appear to account for the great majority of such thermostats (although this is hard to judge based on the available data). Further, as the SERA report points out, flow/disposal of mercury thermostats from large institutional facilities may be more likely to be addressed in a responsible way than flow/disposal from residences. Finally, the draft regulations allow for adjustment of the estimated numbers of mercury thermostats becoming waste each year and performance requirements. Such adjustments, based on actual experience in the field, could be an effective way to address the limitations. For these reasons, while one can identify room for improvement in some aspects of the approach and methodology, overall the results may be sufficiently reliable to support the regulation and proceed with program implementation.

1. **Scientific conclusion to be addressed: Sample Methodology: SERA's approach is reasonable and appropriate and provided representative data on the numbers, ages and types of thermostats in California's households and businesses.**

Comments:

In general, the sampling methodology appears to be based on sound methods and practices. Selecting a simple random sample of households and a stratified random sample of businesses, stratified on size, are reasonable approaches for obtaining representative data on these units. The sampling plan shows an awareness that the number of thermostats should be disproportionately higher at large employers as compared with small employers, and this is addressed by oversampling large employers. The target numbers to be sampled are calculated correctly to provide the desired level of accuracy.

There are some concerns with the sampling frame, however. It is unclear whether non-residential entities that are not commercial, such as schools, universities, libraries, airports, hospitals, utilities, etc., were included in the sampling frame of "commercial" entities. If not, then the inventory and estimated flow are likely to be underestimated.

Furthermore, the questionnaire is not well designed to obtain information from a business involving complex facilities or multiple buildings. Indeed, it would be difficult to construct such a questionnaire or in general to obtain good quality information using this type of mailed/web-based survey for such businesses; the respondents are unlikely to be willing or able to provide accurate data covering the entire complex. For this reason, survey data collected from respondents at such entities are likely to be of lower quality. The SERA

report hints at this; they cite a respondent who gave the number of thermostats as "We have thousands and thousands." Alternative approaches, such as site visits or multistage sampling (e.g., respondent picks one building within the complex and answers with regard to that building) could have provided higher quality data; however, such approaches would generally require more resources.

The relevant question is, are the data obtained "good enough"? Based on ballpark inspection of numbers, it appears that residences account for the great majority of mercury thermostats, which lessens concerns about underestimation and poorer quality data from non-residential buildings. Also, the SERA report does remark, perhaps correctly, that flow or disposal of mercury thermostats from large institutional facilities may be more likely to be addressed in a responsible way, which also lessens concern about this sector.

2. **Scientific conclusion to be addressed: Survey Response Rate and Bias Reduction Analysis: a) the number of responses from California households and businesses was sufficient to obtain accurate data with 95 percent confidence; and b) SERA's bias reduction analysis and correction are reasonable and appropriate.**

Comments:

The response rates for the survey (in general, calculated as number of responses/number contacted) are not provided explicitly. If, as stated on page 10, SERA send postcards to 12,000 business and 10,000 households, then the survey response rates were 267/12,000 = 2.2% and 595/10,000 = 6.0%, for business and households, respectively. The response rates to key thermostat questions were 195/12,000 = 1.6% and 549/10,000 = 5.5%, respectively. These are low response rates, especially for businesses. Response rates on surveys are generally low in this day and age, so this is not entirely a fault of the study; SERA in fact showed good judgment in "re-framing" the survey to be about recycling rather than thermostats to increase the response rate during the re-mailing. However, this low response rate is a concern for both precision and bias of the estimates of thermostat counts.

The number of responses to key thermostat questions from commercial respondents was only 195, well below the goal of 270-380. Thus the margin of error is larger than called for in the study design, at ±7% rather than ±5%, for 95% confidence. The resulting lack of precision is reflected in the point estimate and 95% confidence interval for average thermostats per business in Table 1.1, which are 5.6 and (-0.22 to 11.39). This is a very wide interval (and even includes a negative number!). However, the weighted average of combined commercial/residential has a confidence interval of reasonable width: point

estimate 1.5, 95% confidence interval (0.44 to 2.6). Thus, considering the survey as a whole, the accuracy is reasonable.

Non-random patterns in survey response (whether survey was returned) or item response (whether item on survey was answered) are potential sources of bias. The SERA report addresses the issue of bias in a defensible manner, by calculating and using weights that align the sample with population proportions based on employer size (commercial) or household type (residential). However, there appear to be errors in Table 4-2, Table of Completions and Ratios for Commercial Sector Respondents. (Attachment 5) The counts of respondents by size sum to 187 rather than 267 (or even 258, the number who answered question on number of employees). Also, the response percentages sum to 73% rather than 100%. This should be corrected.

The quality of the study and its findings are definitely enhanced by the validation studies conducted to assess the accuracy of self-reported data on the surveys, although the sample sizes for the validation studies were small and, in the case of the site visits, geographically limited. Adjustment of the counts based on the validation study is a strength of the study, and it is reasonable to use estimated based on these adjustments. However, only 5 of the 30 validation site visits were to businesses. SERA found an 18% undercount based on these 5 businesses, which indicates the importance of such validation data. Again, for the camera validation, SERA sent out cameras to only 7 commercial entities but 37 residential entities. More resources should have been focused on commercial as opposed to residential sites, since these sites are more highly variable. However, it is better to have done some validation work than to have done none.

3. **Scientific conclusion to be addressed: Developing a Method for Estimating Waste Flows: Using the survey approach to obtain statistically valid data and the methodology for estimating the number of thermostats becoming waste annually is a reasonable approach.**

Comments:

SERA used the survey data to construct an estimated cumulative distribution of measure lifetimes for thermostats. This was accomplished using standard survival analysis techniques, using both censored and uncensored (observed) lifetimes. One issue with their analysis is that they used statistical methods that assume all observations are independent, when in fact the data set is not comprised of independent observations on thermostats. Thermostats in the same building are expected to be correlated, i.e., they tend to be the same or a similar type and to have been installed and removed at the same or similar times.

The main consequence of an erroneous assumption of independence for correlated data is that the precision of the estimated survival curve is overstated. There could potentially also be some bias introduced by this assumption, since the multiple thermostats from a single building are not downweighted as they would be if the correlation of observations were taken into account. However, it is reassuring that the estimated distribution of measure lifetimes has a median and maximum that roughly correspond to values from independent credible sources (footnote 26, page 18). This suggests that the independence assumption did not introduce gross errors.

SERA then conducted computations of annual flow of thermostats due to removal from service by combining the distribution of ages of thermostats in the inventory with the expected lifetime model, operationalizing this computation by dividing the inventory into cohorts and obtaining annual disposal streams for each cohort, them summing over the cohorts. They conducted sensitivity analyses, using different portions of the sample and different assumptions, and obtained reasonably similar results. These procedures appear to be reasonable and sound.