Assessing Thermal Stress Benefits in the DOE's Evaluation of the Weatherization Assistance Program

Meredith Fowlie, Michael Greenstone, and Catherine Wolfram[1]

This document provides a more detailed discussion of the issues we identify in the blog. We have examined all of the analysis that underpins the non-energy benefit valuation and we find pervasive problems across all non-energy benefit categories. Here, we examine in detail one particularly important source of estimated benefits: reductions in thermal stress.

The health benefits attributed to changes in thermal stress rely on occupant survey responses rather than direct measures of health outcomes, use an incorrect statistical technique to measure the benefits, and fail to test for statistical significance. Further, there is little change in indoor temperatures, undermining the basis for health benefits associated with thermal stress.

A. Reliance on Survey Questions.

The thermal stress calculations are described on pp. 81-88 of this report. The benefits attributed to reductions in thermal stress are split into two categories: benefits from heat(cold)-related illness reduction, and reduced thermal-stress-related mortality. Notably, the reduction in illness is derived from questions in the national occupant survey: "*In the past 12 months, has anyone in the household needed medical attention because your home was too cold (hot)?*" The survey does not ask questions about mortality, and no other WAP-related data collected by the evaluation team measures mortality directly. Instead, mortality benefits are calculated by converting the question about medical attention to WAP-driven reduction in hospitalizations and applying an estimate of the percent of thermal-stress related hospitalizations that resulted in death.

The exact procedure for converting these survey-based measures into monetary benefits associated with reduced medical care is a long and winding road. There are three main steps:

1. The report calculates the reduction in the medical care needs due to heat- and cold-related illnesses using the new statistical method described below. This reduction purports to measure the change in the percent of households requiring thermal-stress-related medical care, and uses the survey question directly.

2. Medical care is split into three types: hospitalizations, ED visits, and physician office visits. Survey respondents are not asked about these types of care individually; rather, the evaluation uses data from the Department of Health and Human Services (DHHS) to determine the proportions of thermal-stress related medical care that result in hospitalizations, ED visits, and physician office visits.

3. Estimates of household (out of pocket) and societal (insurance) costs for each type of visit are drawn from DHHS. The total benefit is computed as:

[1] We thank Fiona Burlig for truly outstanding research assistance.

Benefits = [(# of weatherized units completed) * (decreased rate of seeking medical care) * (% type of medical treatment) * (average total out-of-pocket medical costs paid by households)] + [(# of weatherized units completed) * (decreased rate of seeking medical care) * (% type of medical treatment) * (average total out-of-pocket medical costs paid by insurance)]

The calculation for converting the survey responses into estimates of the number of fatalities is similarly long:

1. The report takes the number of avoided hospitalizations from step 2.

2. It multiplies the number of avoided hospitalizations by the DHHS-supplied percent of temperature-related hospitalizations that resulted in deaths.

3. It then multiplies the number of lives-saved by the value of a statistical life, $7.5 million. The overall benefit calculation is:

Mortality Benefits = [(# of weatherized units completed) * (decreased rate of seeking medical care) * (% hospitalizations) *(% of hospitalizations resulting in deaths) * (value of a life)]

It is evident that many significant assumptions are required to translate responses to survey questions into monetized measures of medical care and fatalities. These are bold calculations and, even if one sets aside concerns about plausibility, they introduce several sources of uncertainty that are not accounted for in the subsequent calculations.

B. The Invention of a New Statistical Technique.

As described above, the benefit calculations in this section of the report rely heavily on the estimated change in the proportion of households reporting seeking medical care as a result of heat- or cold-related illnesses on the national occupant survey.

We have critical concerns with the way that the "decreased rate of seeking medical care" is calculated. The estimate of weatherization's impact is calculated using survey response from the treatment group (houses weatherized in PY 2011) and the comparison group (houses weatherized in PY 2010) to calculate:[2]

Reduction = [(Pre-treatment – Post-treatment) + (Pre-treatment – Comparison-group 1-year post weatherization)] / 2

[2] It is important to note here that "Pre-treatment" and "Comparison-group 1-year post weatherization" are measured at the same time, but are different groups of households. "Post-treatment" is the same set of households (sans those who have attrited) as the "pre-treatment" group, but measured one year to 18 months later.

This calculation is not a standard method for estimating treatment effects. It makes no attempt to control for the effect of other confounding factors that change over time and almost certainly affect outcomes of interest. To the best of our knowledge, the DOE approach has never been used in any textbook or research paper previously. Additionally, DOE researchers provide no rationale for it recovering the effect of energy efficiency investments.

C. Minor Change in Indoor Temperatures.

The DOE estimates $4,780 per weatherized household of benefits due to reductions in thermal stress. It is natural to assume that the starting point for such benefits would be a substantial change in indoor temperatures, making it warmer in the winter and cooler in the summer. In the only report where indoor temperature is directly measured, DOE finds that average indoor temperature increased by 0.3 F in weatherized homes as compared to a control group.[3] This very small change, calls into question the basis for the claimed health effects due to reductions in thermal stress.

D. The Report Does Not Test for the Statistical Significance of the Key Estimates

The report finds reductions of medical care needs due to cold-related illness of 1.4%, and reductions due to heat-related illness of 1.1%. In neither case is there a discussion of whether these estimated impacts are statistically different from zero.[4] Beyond the analysis of survey responses, no attempt is made to account for the uncertainty introduced by the aforementioned chain of assumptions required to convert survey responses to monetized benefits.

The bottom line is that the report does not shed light on whether the estimated benefits associated with reductions in thermal stress are statistically significantly different from zero. Put another way, readers are left to guess as to whether the reported results are due to random chance.

[3] There is a question in the national occupant survey that addresses indoor temperature: "*How often kept home at unsafe or unhealthy temperature the past year (1=almost every month, 4=never)*". This question does not ask occupants to directly report indoor temperatures. Because of the categorical nature of the question, responses are very hard to interpret. The report does indicate a small increase in the mean response in the post-weatherization period though it is not clear that the responses have been adjusted for outdoor temperatures or other determinants of the outcome.

[4] The only discussion of statistical significance is in footnote 91 on p. 82, where p-values are reported for pair-wise comparisons of surveyed groups but not for the overall estimate of the program's impact. No reported pair-wise comparison between any of the surveyed groups is statistically significant at the 5% level; the difference in means between the pre-treatment and post-treatment cold-related illnesses is statistically significant at the 10% level, as is the difference between the treatment group and the comparison group in the pre-period. There is no discussion of how these p-values were estimated; in particular, no mention is made of clustering to correct standard errors for serial correlation, or of adjustments necessary to account for testing multiple hypotheses. As a result, even if a p-value on the ultimate estimator were to be reported, we are concerned that it would be based on standard errors which are biased towards zero.