

# 2005 National Evaluation of the Weatherization Assistance Program Expert Planning Committee Conference Call May 23, 2005

## Participants

### *EXPERTS*

Sharyn Barata  
Michael Blasnik  
David Carroll  
Greg Dalhoff  
Meg Fels  
Jeff Haberl  
Larry Kinney  
Michelle Marean  
Lori Megdal  
Scott Pigg  
Jane Peters  
Lisa Skumatz

### *ORNL*

Meg Eady  
Joel Eisenberg  
Rick Schmoyer  
Marty Schweitzer  
Mark Ternes  
Bruce Tonn

### *FACILITATOR*

Cynthia Simonson

## Single-Family

### Single Family – Question 1

There was large attrition experienced in the 1990 National Evaluation (about 66%) because utilities did not cooperate in providing billing data or the billing data were incomplete. Considering changes over the last 15 years in utility restructuring and electronic data storage, is this level of attrition to be expected again based on your recent evaluation experiences? What steps have proven successful in gaining the cooperation of utilities? Is the time difference between when the data request is made and the earliest date of data needed important? If so, should data be requested twice (once for pre data and again for post data) to reduce this time difference even if it imposes additional burdens on the utilities? How much over sampling is needed to account for attrition?

- < It is very important to make multiple requests for the data
- < Mergers of utilities can have significant impacts. Before, you lose one utility, you lose one specific area. Now, with the mergers, you lose one utility, and you could lose whole states/regions – a much greater geographic area.
- < Suggest having the clients sign waivers at the outset of Weatherization work, giving the utility “permission” to pass along the data upon request without having to go back to the client for the waiver.
- < Suggest ORNL work through intermediaries for the data. Many have auto download of the Weatherization information and it is often done on a routine basis!
- < Supplemental funding sources may result in a more comprehensive program.

- < Go in and get the data twice – pre-data, as soon as possible (not only because you get it but also) so you can make sure the data is accurate (and useable!).
- < For some utilities, the costs of extracting the data can be as expensive as the full evaluation. Suggest finding out what reports they already have and capitalize on those reports.
- < Once you get the data, throw out houses by screening the data.
- < Grab ALL the data!
- < Don't use data from utilities that have just merged. Their systems are often difficult to work through.
- < Check to make sure your weather data feeds are not changing! When evaluating over a long period of time, this becomes more of a concern.
- < Get the data requests into the utilities as soon as possible. The earlier they have the request, they can start gearing their systems toward providing the information.
- < Getting pre-data should be a part of any good Weatherization Program. In addition to the utility data, other information (blower door readings, etc.) would be useful information.
- < NOTE: It may already be too late for PY 2005 because this is late to start collecting pre-data.

#### Single Family – Question 2

In using PRISM to evaluate the heating energy savings impact of the Program in Texas, both the State and ORNL encountered difficulties in using PRISM to estimate the savings in gas heated houses—the model failed in 30% of the homes. In the 1990 National Evaluation, the variability in savings in gas heated houses in the hot climate region was so great (and the sample size was so small) that a statistically significant reduction (at 0.001 confidence level) in energy consumption could not be confirmed (hot climate region – 590 homes, 102 ccf/year savings, standard error of 24 ccf/year; cold climate region – 1040 homes, 166 ccf/year savings, standard error of 18 ccf/year; moderate climate region – 2243 homes, 137 ccf/year savings, standard error of 15 ccf/year). Have you experienced similar problems in using PRISM to evaluate heating savings in homes in hot climates (i.e., homes with lower heating loads and/or heating loads that are less temperature dependent)? Does such large attrition from model failures concern you considering sample bias or other reasons? Is there a way to reduce the rate of model failures in such homes? How has PRISM been more successfully applied to evaluate heating savings in hot climate homes? How should PRISM be better applied so that a better understanding of energy savings in hot climates results from the evaluation? Should homes in the south be over sampled and, if so, by how much?

- < Suggest a “flatness index” for those parts of the country that has small heating use (south) or small cooling use (north).
- < By doing a full regression analysis, it allows more “power” to do what works in different areas.
- < The 95 version of PRISM does allow for the “flatness index” and was developed for this purpose.
- < (Note – Michelle Marean will check to see if anyone is using the flatness index.)
- < By making “simple degree day” adjustments, summing up seasonal usage, etc., may help make the model work just fine. You can adjust without throwing out the model. Cooling degree days overshoots weather adjustments.

- < Minimal screening – go back and look only at the ones that are the outliers!
- < PRISM is a good screening tool – using PRISM to determine what houses to weatherize!
- < Using “number of days in a row” worked well in one model (especially in the shoulder months).
- < Compare apples to apples...
- < Must have good comparison groups. If the comparison groups work, then we’ll be in decent shape overall.

### Single Family – Question 3

In using PRISM to evaluate the cooling energy savings impact of the Program in Texas, the State and ORNL found that cooling models failed in more than 60% of the homes. Is this experience consistent with your recent evaluation work? Can PRISM estimates of air conditioning energy use and savings replace the need for a submetering study and, if so, how (e.g., should NAC or NACC be used)? If a submetering study is performed, is there any benefit to analyzing electricity data collected on the larger sample of houses used in the heating study to supplement the submetered study?

- < (Answers and comments to this question are already captured under Question 2)
- < Is the submetering sample representative of the larger group?
- < Doing a very careful, well-defined balance point will result in better results.

### Single Family – Question 4

If PRISM models of electricity use fail in 60% of homes heated by natural gas (see Item #3), then the failure rate in all electric homes is likely higher. In the 1990 National Evaluation, application of the PRISM heating-only, cooling-only, and heating-andcooling electricity models seemed problematic (e.g., houses with a heating and cooling peak eventually were analyzed with just the heating-only or cooling-only models). Does PRISM really do a good job with all-electric houses, especially if there is air conditioning present? How can the application of PRISM to all-electric houses be improved? Given that houses heated with electricity are more predominant in mobile homes and in homes in the hot climate region, how should PRISM be better applied so that a better understanding of energy savings in mobile homes and hot climates results from the evaluation? Should electrically-heated mobile homes or homes in the south be over sampled and, if so, by how much?

- < PRISM can run multi-models using one data file. The way the software is currently configured, it can choose the correct model and increases the reliability.
- < Get the data, based on what the end uses are.
- < Getting a representative sampling of buildings is the most important. Relax the screening requirements.
- < Behavior factors. E.g., a client is behind on a gas bill so the client starts using electric baseload.
- < Observation – In all income brackets, what the utility thinks is most likely NOT happening in the home. Let the data decide what is happening.
- < Alternate Observation – Sometimes the data reported is not accurate (e.g., data shows up as a cooling signal when there is no real cooling).

### Single Family – Question 5

Billing data and PRISM cannot generally be used to evaluate homes heated by bulk (unmetered) fuels. Two large groups of such houses are houses in the northeast heated by fuel oil and mobile homes heated by either propane or oil nationally. Is submetering the only analysis approach? If yes, what metering/analysis approaches (e.g., Minnesota's ASAP and DESLog, NREL's PSTAR and STEM short-term methods, Energy Research Center's and John Batey's bulk fuel meter) are new, have been used recently, and/or might be used in a submetered field test?

- < PRISM can model oil, propane, or other fuel if the data is there.
- < You have to submeter! Sometimes the population gets half a tank so it is difficult to get reliable delivery information.
- < 2001 RECS collected fuel oil data.
- < PRISM study for NYSERDA estimated readings could be used for run time data. Suggest PRISM be used as the starting point.

### Single Family – Question 6

If a submetering study is needed to determine air conditioning energy savings (see Item 3), what metering/analysis approaches (e.g., ORNL's use of NIALMs in Texas, NREL's PSTAR and STEM short-term methods) are new, have been used recently, and/or might be used in a submetered field test?

- < Look at billing data in lots of houses!
- < Use the most simple stratification sampling plan.

### Single Family – Question 7

Can PRISM be used to measure baseload energy savings (principally from hot water, refrigerator, and lighting retrofits) or is a submetered study needed? Can PRISM estimates of baseload energy use and savings supplement submetering? If submetering is needed, what metering/analysis approaches (e.g., ORNL's use of NIALMs in Texas) are new, have been used recently, and/or might be used in a submetered field test? How does the fact that PRISM analyses of houses heated by natural gas and electricity often already include baseload savings (hot water for both gas and electricity, refrigerator and lighting for electricity) complicate the determination of baseload savings and developing a national estimate?

- < PRISM should not be used necessarily.
- < Submetering is expensive and not as accurate as billing data.
- < On the theory that we have analysts that know what they are doing, use large samples and comparison groups!

## Multi-Family

### Multi-Family – Question 1

In the 1990 National Evaluation, the energy savings of many buildings (50%?) could not be analyzed with PRISM while performing the case studies. In buildings with central heating systems using a metered fuel (usually natural gas), what limitations or difficulties are there in using PRISM? How can PRISM be best applied to ensure model success? Is weekly rather than monthly data needed? If yes, how can such data be obtained?

- < Meter NYC (or at least work with the delivery mechanisms).

- < Weekly data does not seem to make a “robust” difference. If you have it, use it, but results are not worth additional effort to collect.
- < Rental characterization in Wisconsin, where we had data, PRISM worked really well.
- < Adopt the flatness index or you end up dropping out units in the building.
- < Research Question – After Weatherization, do you have a higher occupancy rate, therefore higher utility bills for the building?

#### Multi-Family – Question 2

In buildings with individually metered heating systems (natural gas or electricity), do you know of any limitations or difficulties in using PRISM applied to the units individually (unit-level analysis)? How can PRISM be best applied to analyze the energy bills for individual units? If the whole building was weatherized but the units are individually metered, should a building-level analysis be performed rather than a unit-level analysis by aggregating energy bills? How do you deal with common area measures if the analysis is performed on a unit-level basis?

- < Either way works!
- < The advantage of doing aggregate is being able to do comparison with other buildings. Encourage doing both if you have the data for both.
- < They did request building data but only got the information for 5% of the buildings. You should expect better response this time!

#### Multi-Family – Question 3

How should the evaluation deal with buildings with central heating systems that use a bulk (un-metered) fuel? Is submetering needed? If yes, what metering/analysis approaches (e.g., Minnesota’s ASAP and DESLog, NREL’s PSTAR and STEM shortterm methods, Energy Research Center’s and John Batey’s bulk fuel meter) might be used in a submetered field test? How should buildings that use dual fuels be handled?

- < It is easier to submeter the multi-family buildings.
- < Most of the submetering will be in NYC, which will be easier than having it spread out across the country.
- < DesLOG works fine for single and multi-family buildings.

#### Multi-Family – Question 4

How do you recommend that energy savings analyses done on a building level (e.g., because the building has a central heating systems, is master metered, was weatherized as a building, etc.) be integrated with those done on a unit level (because the unit had its own heating system and was weatherized separately from the building)?

- < This is apples and oranges.
- < Are these really large multi-family? Information collection characterization is important so you know what group the unit falls in, then you can determine how to study it.
- < Weatherization definition of multi-family is a building with 5 or more units. Do we deviate from that definition for the evaluation?
- < Apply the definitions consistently.
- < Number of stories is more important than number of units.
- < However, size affects central units and how they perform.

### Multi-Family – Question 5

What comments would you have on a sample design that focused on multifamily buildings weatherized in New York City (50% of total in 1990), California (25%), and then others?

- < Based on the characterization, it is important to have good sampling. Get a good cross-section, not just the very high rises of NYC and Philadelphia.
- < Gathering data that would be useful in improving the program would be a great output of the evaluation.

### Multi-Family – Question 6

How should control groups be selected?

- < Get the consumption of all LIHEAP applicants (applicable for both single and multi)
- < Use only people that applied in the last three months.
- < Pipeline buildings.
- < State LIHEAP offices may not share the data.
- < If time is an issue, match them up by time period. Look at those treated early in the year with a comparison group that was treated early in the next year so it balances out.
- < Get good data on treatment dates. Any error on treatment dates makes the savings look lower.
- < Often, we only have completion dates, not the start dates.

## Non-Energy Benefits

### Non-Energy Benefits – Question 1

Should the study of non-energy benefits utilize averages from current literature or generate new primary data? What are the pros and cons of each of these general approaches? Should *both* approaches be used?

- < Both, but, some can piggyback on data better than others
- < Depends on what benefit category you are talking about!! In broad terms, if you get info from utilities themselves, you could get the impacts.
- < Need to determine what benefit categories you are going after – the objective is to quantify the benefits the same way we would quantify savings benefits.
- < Financial transaction data comes in so many formats, it doesn't seem like a viable thing to go after. These vary dramatically by utility.
- < Meter reading codes that show shut off and reconnections. This becomes part of your usage data.
- < Look at fuel assistance payments almost always drop off in the year following Weatherization.
- < Mobility impacts – high estimates. Look at thorough utility billing data. There may be waiver issues there too.
- < Review papers and have the analysts weigh in on what should be looked at.
- < Building analysis and building models on average buildings – calibrated pre- and post- and did a qualitative study on comfort. How much did we close the gap in the temperature differential? It can be quantifiably measured!

- < Process evaluation should cover comfort, security, etc.
- < Look at what the Program is meant to achieve!!

#### Non-Energy Benefits – Question 2

At what geographic level should the non-energy benefits estimates be made? Nationwide? By climatic zone or other multi-state regions? State by state?

- < Climate zone level – depending on how uniform the program is run.
- < Water savings? Full range of tests?
- < Is there a reason to get down to the state level information? If you want to prod a state to move forward in a specific area, then state-by-state information is helpful. However, you may not be able to get the data on a state-by-state basis.

#### Non-Energy Benefits – Question 3

Are updated data available on the number of jobs created per weatherization dollar? What are the pros and cons of the methods available to collect additional data on this topic?

- < Just impacts on the state or local economy?
- < Input/output modeling works well.
- < Boundary issues – you can show jobs created but there isn't an evaluation of jobs that were NOT created.
- < Would the money have been spent elsewhere by the government?
- < Several have looked at the net basis and alternative uses for the money. This money "could have been used in LIHEAP rather than WAP..."
- < What would be spent in the country versus outside the country?
- < Power may be shipped in from somewhere else – state-by-state may vary dramatically.

#### Non-Energy Benefits – Question 4

What can be done to reduce the huge range of values associated with current estimates of environmental benefits? What are the relative merits of the "Market Valuation" method and the more comprehensive approach that attempts to include all benefits to human health and ecosystems?

- < Market valuation is the method that should be used when there is a market for an emission.
- < SIP credits – demonstrating to EPA means counting at the county level and making multi-year projections. To reduce the range is to demonstrate the method (17 regions within the state).
- < Should include the health benefits under one topic or the other.
- < Greenhouse gas will be the "wild card." It is one of the largest percentages.

#### Non-Energy Benefits – Question 5

What are the best methods available for quantifying national security benefits, and what are their major strengths and weaknesses?

- < Very little information in this category. Very speculative and very political.
- < May not even want to go there!

### Non-Energy Benefits – Question 6

If primary data are to be collected at individual weatherized houses, what data are most appropriate to gather (e.g., CO measurements, indoor temperature)? For what size sample?

- < Gas check information was very valuable in the Ohio study (info the agencies collected and the CO levels and gas leaks on one or more appliances, and high draft). How many of those were addressed by the program?
- < Indoor temperatures – piggyback on a metering study.
- < CO<sup>2</sup> is a good indicator of tightness. Temperature and humidity – regression against a delta T rather than temperature. On a daily level, that works well.
- < South – good temperature control but difficult humidity control. Detect duct leaks, diagnostics measures.
- < Comfort Partners Program in NJ – good job of detecting gas leaks but to do a good job, you have to go into the individual records at the agency level. “Major gas leak” is in the notes of the audit, but not in the H/S portion of the audit.
- < Often, the problems aren’t fully resolved.

### Non-Energy Benefits – Question 7

Should individual participants be surveyed regarding their responses and reactions to the weatherization? If so, how many should be surveyed? How should the sample be stratified? Should a control group of non-participants also be contacted?

- < Often relate back to the goals of the program, so yes, survey.
- < Net of positive and negative impact; net of gross.
- < Depending – something efficient versus non-efficient. Only measure the difference (in the event someone was going to do some type of “weatherizing” anyway.
- < Would they have bought a new appliance or gone to a “used appliance store.”
- < Pretty focused, because doing a survey is often thrown out because of the expense.
- < Occupant surveys – In Habitat Houses (with forced air systems) – the heat exchanger fails at twice the rate as the manufacturer suggests. How can we design houses (especially in the south) that don’t have that problem?

## **Explanatory Factors and Process Variables**

### Explanatory Factors and Process Variables – Question 1

What are the central issues associated with evaluating client education programs? How do you propose designing such an evaluation? How many states and agencies should be included in this study?

- < It can be done. It just costs a lot of money to get this.
- < Education is only measurable if it is very specific and targeted.
- < Get the educational materials and have a professional group write those materials. What do you get from the additional money or higher quality materials? Maybe a less expensive way to find out what you get from education.
- < Program regression models – this has “good” education. If they have good education materials, they probably have good everything else too! The better savings may be more related to measure installation.

- < Study idea – is there an indication when people move out, the client ed goes with it. Large sample would be needed to conduct this study.
- < What is the value to the participant? Low-income households often feel powerless. Making that choice with information is giving them control. That is worth something. Look at evaluation design with savings and then power.
- < Recognize these consumers are different than other consumers. This realization can be useful. And are they “educated” or are they just handed a brochure?
- < If you do look at “control over the bill” because of education, it has a high impact on participants.
- < State at the state level to see what the quality of the client ed is currently. It may be helpful to get that and leverage in “controlled experiments” as a sidebar.

### Explanatory Factors and Process Variables – Question 2

What are the central issues associated with evaluating weatherization training? What are the evaluation questions? How do you propose designing such an evaluation?

- < Different “spins” in the way states do training. Variety of different types of training that is happening. This study should document the range. In general, training programs are very specific.
- < All states do training, but what works and what should each state be doing?
- < Best practices would be helpful here!
- < DOE’s goal, it is isn’t to have all programs look the same, then 1) pull out “best practices” from what people tell you have been successful and 2) find best programs and go look at how they do things.
- < The process? How has it worked from the national level and how has it been accepted? How could the training be improved?
- < How do you determine what state is more effective? “10 high performing agencies” – Much more the strategy of the agency employed for getting savings. It is difficult to correlate that with training!
- < Direct observation – double blind and see how performance correlates to training.
- < Sample of field personnel and give them a test!
- < Inter-agency savings is often driven by the houses they treat.
- < Crew versus contractor staffing makes a big difference in the quality control mechanisms.

### Explanatory Factors and Process Variables – Question 3

What explanatory factors and process variables affecting energy savings and cost effectiveness should be examined, what data are needed to perform analyses, and how should these data be analyzed to determine which are important and their relative rank or weight? Is an analysis of high and low saving houses and/or agencies useful?

- < Must sample across the state and agencies so you actually have the ability to draw some conclusions that are cross-cutting.
- < Ability to follow a prioritization scheme. Bulk fuels are not as successful as electricity and natural gas based on pre-usage.
- < Obvious predictors – wall insulation, heating system replacement, air leakage reductions, attic insulation – how many measures were done and how much money

was spent? Often the variations between agencies is because of the housing stock (mobile homes = lower savings).

- < Are the better agencies doing wall insulation more often OR is it because they have a lot of wood framed houses without any insulation?
- < If what agencies need to know is “how can we do our jobs better?”, then concentrate on answering that question. How much difference does it make “whatever the things are they can make decisions on).

#### Explanatory Factors and Process Variables – Question 4

Should we conduct in-depth, comprehensive case studies? If yes, what energy savings, process area, building type, etc. should they focus on? For example, we could choose a few states and interview state-level people who administer the program, numerous agency staff, numerous weatherization crews, appropriate utility staff, and weatherization recipients to create entire pictures of a few programs. Is this worthwhile? If so, do you have any advice about how to design the case studies?

- < Where are some good things that have happened and why did they happen? What can we do to encourage these good things to be replicated?
- < Case studies are great to put a human face on a human study. These are the things that cut through the statistics.
- < Case studies work much better when looking at technology but there is so much diversity in the program, getting “snippets” can be confusing.
- < Case studies of agencies are valuable to understand how these agencies relate to the community.
- < Case studies are not for the purpose of evaluating.
- < A “process evaluation” is not a statistical document, but does give a human face – but at a level that is much more comprehensive than a case study.
- < There is a lot of room for abuse with case studies.
- < The “case” for case studies is if you want some understanding behind why the numbers vary by state or by agency. That may come as part of the process and impact evaluations, but, we may find some areas where a case study will then explain some of the variances and give a face for some instances.
- < Unless there is a specific purpose for the case studies, a process evaluation is much more valuable. Asking how to improve the program, opportunities for improvement, etc.

#### Next Steps

- < Meeting summary and white board notes are sent out to the committee.
- < Committee provides additional thoughts/ideas/inputs to ORNL. ORNL will incorporate and send back out to the committee.
- < ORNL team will begin drafting which elements of the overall evaluation process will be going out for competitive bid. (Note – ORNL may ask for your thoughts on elements that will NOT go out as a competitive bid but anything that may go out in a RFP will not be sent to the committee for input).

**How did this process work (using web-based meeting software) as a tool for conducting this conference call?**

- < Worked pretty well
- < Handraising worked well.
- < Kept us on schedule.
- < Would like the ability to post attachments.
- < People still need to identify themselves since it often was difficult to know who was commenting.
- < Make sure people know this software doesn't work in Netscape.