Seven Years After: Impact Evaluation Results Employing Extensive Site Inspection Data and Associated Pre/Post Billing Analysis

Anne Minor West, Anne Minor West & Co.
Howard S. Reichmuth, MicroGrid, Inc.
Pamela Brandis, Bonneville Power Administration
H. Gil Peach, H. Gil Peach & Associates

This paper presents a longitudinal impact evaluation of a 1986 residential retrofit program at a Northwest municipal utility. Extensive site inspections and measurements of both participant and non-participant residences assayed the air tightness of the buildings, the showerhead flow rates, and (for participants only) the quality of retrofit measures seven years after retrofit. Utility records were reviewed for the mix of measures and utility personnel were interviewed regarding policies extant at the time of the program. An analysis of ten years of electric billing data for each building provides a pre-retrofit baseline based on three years of billing data, compared to two successive three-year groupings of post-retrofit billing data.

Billing data aggregated into three-year groups is manually edited based on site occupancy audit data, and a single linear break-point function is least squares fitted to the edited data. All aggregate statistics are building area-normalized to rationalize area differences between the participant and comparison groups.

An estimate of net program savings is made by including analysis of billing results for a non-participant control group. Net savings are clearly evident for an identifiable subset of participants referred to here as ''probable savers.''

Methods used here were applied to a much larger 1988 and 1989 data set, and results are as yet unpublished.

INTRODUCTION

This study is an assessment of the persistence of energy savings achieved through the Bonneville Residential Weatherization Program, (RWP).[1] As the host utility in this study, the Eugene Water and Electric Board (EWEB) provided access to data and housing stock in their service territory. The target of the study is the physical impact of the program over a period of six years following the retrofit. The physical impact of the program is assessed in terms of changes in billed energy use and also in terms of site interviews, inspections, and measurements of the current condition of the retrofit measures. The study is aided by a review of previous program evaluations of RWP,[2] and by a review of EWEB program documentation.[3]

Historically, the overall goal of the RWP was to use the economic engine of resource acquisition to improve the energy efficiency of the existing residential housing stock. The direct savings proceeding from the program retrofits were an important consequence of the program, but it was also hoped that the program would contribute to a ''market transformation'' in energy conservation services and products which would have a broad regional impact on residential energy efficiency. Program policies favored inclusion of participants with significant wood use as part of the regional consensus that supported the program.

Potentially, the net savings from the program could legitimately be quite complex, including electricity and wood usage changes occurring directly by retrofits and indirectly by market transformation, takeback, free riders and free drivers. Fuel switching adds another dimension to an assessment. However, the heart of even the most complex savings assessment is the physical event of the retrofit and the durability of the retrofit measures.

In contrast to other impact evaluations of this program, which use large samples and no site visits, this study is much more restricted. The emphasis here is to use extensive site observations, field verification, and diagnostic testing, coordinated with analysis of ten years of billing data as acquisition validation techniques to identify the thermal building changes due to the retrofit and to assess the longevity of the retrofit measures.

A sample of 20 homes participating in the 1986 Bonneville sponsored Regional Weatherization Program were investigated. The participant sample was chosen from homes originally submitted to Bonneville for analysis through the Data Gathering Project.[4] A second sample, a non-participant comparison group of 16 audited and unweatherized homes, was also investigated.

SUMMARY OF RESULTS

The results of this study are instructive to future residential retrofit program design. (1) The condition of the measures after seven years was generally good, and net electrical energy savings persist essentially undiminished for the six post-retrofit years examined. (2) The participant group divides into two categories, probable savers and probable non-savers. The probable savers could have been identified prior to the retrofit by using area-normalized billing versus temperature data. (3) The levelized cost of savings for the probable savers was 41 mills/kWh, well within the cost-effectiveness horizon of the program at the time of its execution, and about as expected in the program design. (4) The principal reason for poor savings results was fuel switching, and/or wood use, and take-back associated with increased space utilization or increased use of the electric heating system. Participants with these characteristics were admitted to the program because of eligibility policies intended to be broadly equitable to the utility customer/owners who were ultimately responsible for the utility indebtedness.

METHODOLOGY

Sample selection

A stratified sample, rather than a random sample, was chosen for the participant group in order to be representative of several significantly different physical types of participants. The participant stock was stratified to include representatives of single and multi-family participants and of ducted and unducted electric heating systems. The sample was restricted to participants in the 1986 program to provide a six year interval of measure attrition.

A non-participant comparison group was chosen from the pool of about 14,000 homes which had been audited by EWEB and were not weatherized through an EWEB Program at any time since 1986. The comparison group was restricted to bona fide electrically heated homes in order to reveal more clearly in the billing data, the groups' behavior with respect to electricity price elasticity. The comparison group was also employed to investigate the energy conservation measures in place and the differences between participants and non-participants. The comparison group was approximately matched to the participant group in terms of conditioned area.

Since the total sample consists of 20 participant residences and 16 comparison residences, it will support only limited statistical precision. This sample can reveal evidence of large departures from expectations for the participants and for the comparison group.

The participant sampling criteria were chosen to yield a clean sample which would perform optimally. However, because of limitations in time and circumstance, the criteria were modified. Final selection criteria included: (1) The selection of homes from the Data Gather Project (DGP) providing a link to homes analyzed in the yearly Bonneville project report series. (2) Full year occupancy from 1983 forward for single family homes. Multi-family residences showing more than a one month vacancy between tenants were rejected. (3) Screening for auxiliary heat was attempted, with an allowable limit of 1/2 cord per year.[5] (4) Homes remodeled or which had switched wholly or partially to gas were included. These changes occur normally in the residential population. It was important to determine the impact these changes had on residential savings.

Factors influencing the final sample also included the measures installed, and, the ability to convince owners that the four hour visit to the home would be informative and useful to them, as well as to our research and evaluation efforts. On the whole, those who chose to participate felt that they had benefited from the weatherization work and this was an opportunity for them to do something in return to help the utility, as well as learn more about their homes. No other incentive was offered to gain participation in the study.

As with the participant group, no incentive was offered to comparison group candidates. Owners who chose to participate expressed interest in the updated and extensive audit information about their homes which they would receive, and usually expressed goodwill toward EWEB.

Field verification protocol

EWEB records for each site were carefully analyzed. Records included the billing and consumption histories as well as documents relating to participation in the weatherization program. These included, for example, audit sketches, analyses, measures installed, the data gathering sheets, and job cost data.

On-site inspection and verification procedures were tailored to the objectives of this study. A sixteen page set of Field Verification Forms was developed specifically to meet the needs of this project as it proceeded. Forms were designed to gather data necessary to track demographic, behavioral, and physical changes to the home, develop a baseload profile including short term metering, identify water conservation potential, identify measure installation quality, effectiveness, and persistence, collect test data from blower door, pressure diagnostics, and duct leakage diagnostic procedures, and to document observations made with the infra-red camera. Photo-documentation of installed measures or site conditions was also used to record information. For each home, a ‘‘site book’’ was compiled which included background and field data, analysis results, and technical summaries.

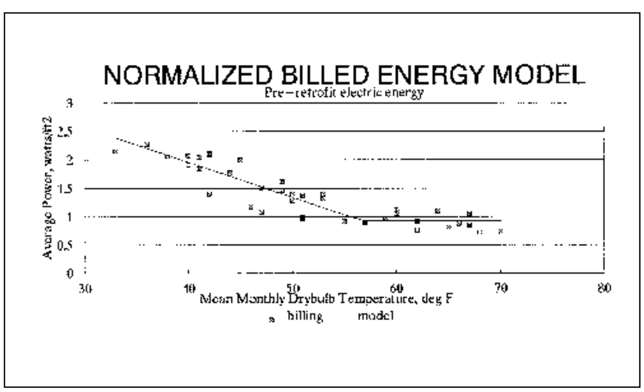
Billing analysis protocol

The billing analysis protocol is intended to derive a physical measurement of building performance changes due to the retrofit. In order to use billing records for a statistically valid physical measurement, temperature data is combined with the billing data and the combined billing/temperature data is further aggregated and normalized. The aggregated and normalized data for a single building is modeled by least squares fitting to the data a simple function describing the building energy use in terms of watts/ft^2. For both the participant group and the comparison group, the modeled data from all sites is further aggregated into a group performance model which describes the performance of the entire group or a particular subgroup in terms of watts/ft^2 vs. temperature.

Successive normalization of the data

Billing data from each of the 36 sites was reduced to a comparable basis so that data from different sites or time periods could be examined together. This methodology normalized for three principle factors. (1) AVERAGE POWER—Billing data was normalized by the number of billing days in the meter read cycle to convert the billed energy to average daily energy in kW or kWh/day. (2) BUILDING SIZE—At each site, data from the billing model was reduced to a ‘‘per square foot’’ basis to compensate for variations in building size. (3) WEATHER—Billing data for each month was associated with the corresponding mean temperature over the billing interval in a manner similar to a PRISM®[6] analysis. The normalized date is presented in a graphic format showing the average building power in watts/ft^2 versus mean temperature as in Figure 1.

Figure 1. *Normalized Billing Data and Model*



Three year aggregation of the data

For this analysis, the ten years of billing and temperature data for 36 sites were divided into three groups of three years each, with the most recent year, 1993, used as an inspection reference. (1) 1982–1985, designated Pre data, describes the pre-retrofit condition. (2) 1986 is not used because pre and post retrofit data are mixed, as this is the year weatherization measures were installed. (3) 1987–1989, designated Post1 data, describes the first three years post retrofit. (4) 1990–1992, designated Post2 data, describes the 4th through 6th years post retrofit. (5) 1993 data describes the performance of the building in the year of this study’s inspection. It provides a check for identifying whether the performance of the building as inspected is representative of its longer term performance.

A fit to three or four years’ data is much more stable and representative of the building’s physical characteristics, the primary targets of this exercise. The consolidation of the data for analysis into these multi-year groups sacrifices some year-to-year resolution in favor of more precision in estimating the physical characteristics of building energy use.

Modeling the normalized annual consumption, NAC, for a single building

The modeling method used for the weather normalization is conceptually similar to the PRISM® method used to normalize large data sets. However, the method used in this analysis differed from PRISM® significantly in terms of the data point selection criteria.[7]

For each site, data in each of the three year-groupings (i.e., 1982–1985, 1987–1989, and 1990–1992) was characterized by a model as described below. The models were used to estimate the average annual energy use for those years for a standard set of monthly average temperatures. By this use of a model, a Normalized Annual Consumption (NAC), is estimated for each building.

This NAC can be used to compare a building’s performance during one multi-year interval with its performance during another. The weather normalized annual savings during the first three years after retrofit are then NACpre—NACpost1 and the savings for the fourth through sixth years after retrofit are NACpre—NACpost2.

In principle, a model could be fitted to each year’s data to derive a building’s NAC for each year. However, the inherent noise in the data is enough to make the year-to-year fits quite variable.

The model fitted to the normalized billing data is a simple linear change point building energy model. For all temperatures above a certain temperature, designated as the ''balance point,'' the energy use in kWh/day is constant. For temperatures less than the balance point, the kWh/day increases linearly as the temperature decreases. The model is fit so that the least squares difference between the model and the data points is minimized. Before fitting the model certain data points are removed from the fit if they appear to represent a significant behavioral change unrelated to the physical characteristics of the building. Figure 1 also shows the fitted model.

The normalized model has four parameters. (1) BUILDING AREA (in square feet)—This is the area within the conditioned envelope. (2) EFFECTIVE BASELOAD (in watts/ft^2)—This is the average daily energy use for all months with average temperatures greater than the balance point temperature. It is physically a measure of all the daily energy uses except space heat. (3) EFFECTIVE UA (in BTU/deg hr/ft^2)—This is the slope of the temperature dependent portion of the model. It applies to all months with temperatures less than the balance point temperature. This slope has been transformed from kWh/degree-day to the English units BTU/degree-hour (BTU/deg hr) to ease comparison with the audit estimated UA estimates expressed in the same English units. This slope is commonly in the range of .5–.8 of the slope calculated as the audit UA for the building, calculated by means of the ASHRAE steady state heat loss methodology. The audit UA is based on the whole building experiencing a steady state temperature difference, while in the real world, usually less than the whole building will experience a cyclically varying temperature difference. Therefore, it is physically reasonable to expect the billing slope to be less than, but correlated to, the calculated UA. The billing based estimate of the UA is called the ''Effective UA'' because it implicitly includes all the effects of real world internal gains and varying thermostat setpoints. (4) BALANCE POINT TEMPERATURE (in degrees F.)—This is physically the lowest average monthly temperature for which the building can be heated by its internal gains alone, without the use of space heat.

Method for estimating building NAC savings

The Normalized Annual Consumption, NAC, is the sum of the monthly consumptions established by evaluating the building model at each of the TMY monthly means. NAC savings refer to the changes in annual energy consumption. Other studies refer to these savings as DNAC (''Delta NAC'' or the change in NAC).[8] For each building, NAC savings are estimated for the years 1987–1992 as follows. (1) 1987–1989—Each of these years is assumed to have the same average savings estimated as NAC Savings = NACpre—NACpost1. (2) 1990–1992—Each of these years is assumed to have the same average savings estimated as NAC Savings = NACpre—NACpost2.

Aggregation into a group performance model

Since the individual building models have been normalized to a form of watts/ft^2 vs. temperature, the models for all the buildings in a group can be aggregated into a single model, identical in form to the single building model in Figure 1, which describes performance of the whole group in terms of watts/ft^2 vs. temperature. This group performance model is the area weighted average of all the individual building performance models, and it is algebraically equivalent to the use of the individual models in computing the savings for the group.

Calculation of net savings using comparison group data

The normalized group performance model provides a basis for estimating net program savings by applying a correction for the changes in energy use by the comparison group. The change in the comparison group performance from the pre to the post retrofit periods is the modeled difference in the control group performance, pre vs. post, in terms of watts/ft^2 vs. temperature.

FIELD OBSERVATION RESULTS

Physical measure persistence and observations

Measure degradation fell primarily into two categories, active degradation caused by people, and, passive degradation which is normal degradation over time.

Insulation degradation. Active degradation of the insulation was caused by people who were in attics or under the floor making repairs, storing items, or remodeling. See Table 1 summarizing measure persistence.

In attics where no one had visited, the insulation remained in good shape. Some settling had occurred. Nearly all baffles were still intact. The few soffit vent baffles that had become detached did not block the soffit vent, but did slightly compress insulation. Underfloor degradation considered ''normal,'' after inspecting the homes in this study, included about a 1″ to 3″ sag in the batts. String or wire support systems fared better than lath. It was not uncommon to see delaminating batts, or separations between adjacent batts where lath spacing was too far apart or the lath had become detached from the joist. Very few homes had large sections in disrepair where there had been no human intervention.

Using the infra-red camera, only two homes were found with insulation voids in a single wall cavity. Insulation coverage was good for its vintage. However, with the infra-red camera it could be seen that leakage at the top plates, and leakage at kneewall/attic interfaces, was common to virtually all homes, and had not been remedied in the retrofit. Homes remodeled since retrofit exhibited leakage at the old and new interfaces.

Thermal boundaries. Several homes visited in this study had ‘‘confused’’ thermal boundaries. Some of the difficulty in determining the thermal boundary arises when homeowners claim to use the space or are about to remodel in order to include the area as living space and then do not follow through. Or, an unconditioned area which cannot be isolated from surrounding living space remains untreated. It can happen then, that as the house behaves under normal operating conditions, the treated unfinished attic or basement continues to function as an exterior zone, and leakage from the house proper occurs. This also occurs when an area fully utilized and treated as living space during the weatherization process is later closed, although not thermally isolated, and is not heated.

Duct degradation. Active degradation of the ducts was caused by people and animals. Passive degradation was caused by poor installation. All six homes with ducted heating systems had significant duct leakage. Supply leakage to the outside ranged from a low of 147 CFM @ 50 Pa to a high of 315 CFM. Duct leakage and its impact upon residential energy use is currently one of the studied and researched topics in the industry. The energy use impacts of duct leakage are unique to each house. Examinations including only changes in the UA of a house will not account for the complex relationship between a forced air heating system and consumed energy of the house.

Window degradation. Storm windows and window replacements were largely in place and in good condition. Few owners reported broken seals in thermal pane windows. Few homes had noticeably loose fitting storm windows, failing caulk, or exhibited workmanship of inferior quality.

Comparison group inspections. Homes in the comparison group had varying levels of insulation. With the exception of one owner who installed thermal pane windows, no owners pursued the installation of weatherization measures on their own. In these comparison group homes, the most outstanding item in terms of potential savings were homes with the combination of ceiling cable heat and uninsulated floors. The ceiling radiant heat was directed at the cold floor. The dynamics between these two elements of the building envelope should be investigated.

Interviews. Extensive interviews with owners which covered demographic, behavioral, baseload, and structural changes were used in conjunction with the billing analyses. Interviews covered the analysis period, from 1983 through 1993.

Field data and associated interviews lead to the following general observations. (1) Non-participation was often cited as due to: cost; other priorities; and, the house is ‘‘good enough.’’ (2) Observable changes in consumption identified in the analysis of heating and baseload data could nearly always be tied to specific events, such as changes in family size, (births, deaths, movers), and installation or removal of additional load (e.g., RV and garage space heaters). (3) Reasons for homes exhibiting poor savings results were fuel switching, and/or wood use, and voluntary take-back associated with increased space utilization or increased use of the heating system. (4) In both the participant and comparison groups, the amount of wood used for supplemental heating was often underestimated by the users. In this study, 2 comparison homes and 6 participant homes used at least one cord of wood per year for supplemental heating. (5) Most homes cannot be categorized prior to weatherization as to which will increase wood use, decrease electricity use, or switch fuels after participation.

BILLING ANALYSES RESULTS

Participants can be divided into ‘‘savers’’ and ‘‘non savers’’

Examination of the models for each of the participant buildings shows that the building temperature sensitivity parameter, effective UA (Btu/deg hr/ft^2), was a key predictor of NAC savings. Figure 2 shows the participant group ranked by Effective UA/ft^2 for the pre-retrofit period. This figure shows that significant changes in building thermal performance occurred for only about one half the participants. Notably, the participants who showed large savings could have been identified in advance, prior to the retrofit. For the purposes of this study a subset of the participants has been selected for separate aggregation and designated as ‘‘probable savers.’’ In Figure 2 it is evident also that the post retrofit thermal performance is about .3 BTU/deg hr/ft^2, approximately what could be expected from a super insulated new home. Therefore, if the thermal performance of the building prior to retrofit is less than .3 BTU/deg hr/ft^2, it is probable that the building has an alternate heating source or is underutilized, both of which will cause the building to be a ‘‘non saver.’’ Buildings with a pre-retrofit effective thermal performance greater than .3 are considered ‘‘probable savers.’’

The performance of the ‘‘probable savers’’ group changed from pre to post retrofit, but the group performance of the whole participant group changed very little from pre to post retrofit

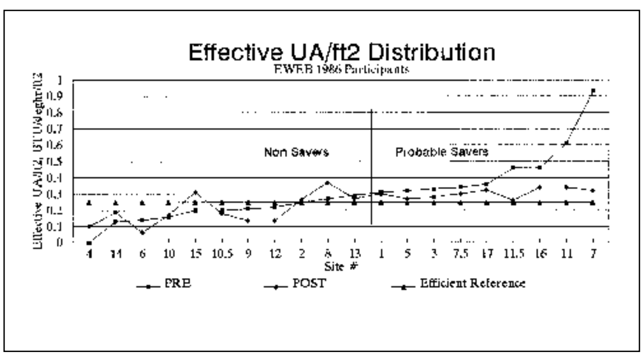
Figure 3 shows the group performance models for the whole participant group and for the ‘‘probable savers’’ subset. It

Table 1. *Conservation Resource Inventory*
Measure persistence 7 years after installation: percent intact

Site	Ceiling Ins.	Floor Ins.	Ground Cover	Pipe Ins.	Windows	Wall Ins.	Duct Ins.	Caulk	Weather-strip
1	95	90	100	98	100			50	50
2	90	95	100	100	100	95		100	
3	100	100	95	100	100	100		100	100
4	95	95	90	90	100			100	100
5	80	80	90	90	23	100			
6	90	95	100	100	100			100	100
7A,B	100				50	90		25	25
8	30	90	100	100	60	60		60	60
9	80	90	100	0	100	95	70	100	
10A,B	90				100	100		60	100
11A,B	95	90	100	100	100	95			0
12	90				100	100		100	100
13	10	80	80	50		60	50		50
14	100	90	100	100	100	100			100
15	98	95	100	100	100			100	100
16	90	90	100	100	100			100	100
17		80	100	50	90		75	100	100

Sites 5, 8, 13, and 16 were remodeled between 1986 and 1993.
Site 7AB received incorrect ceiling insulation application.

Figure 2. *Partition into Savers and Non Savers*



appears in Figure 3 that the effective UA of the probable savers was reduced by the retrofit.

Modeled in this way, the control group shows an increased energy use of up to .25 watts/ft^2 depending on the temperature as shown in Figure 4. This increased energy use in the comparison group is the correction used to derive net savings from the gross savings. The net savings are also shown in Figure 4 for the whole participant group and for the probable savers subset. It is evident in Figure 4 that the net savings for the whole participant group and for the probable savers has persisted undiminished for the six post retrofit years examined. It is also apparent in Figure 4 that the probable savers had more than twice the savings/ft^2 than the whole participant group.

Net levelized cost of savings

A net levelized cost of savings has been estimated for the whole participant group and for the probable savers subgroup. The levelized cost of savings is computed assuming a 30 year measure life, 5% real discount, and net savings decreasing at 5% per year for the unexamined years 7 to 30. The levelized cost and other summary information is presented in Table 2.

Figure 3. Group Performance Model

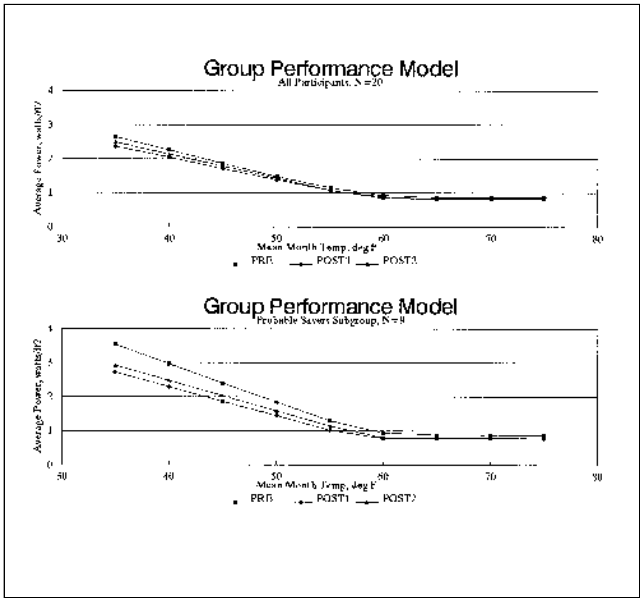
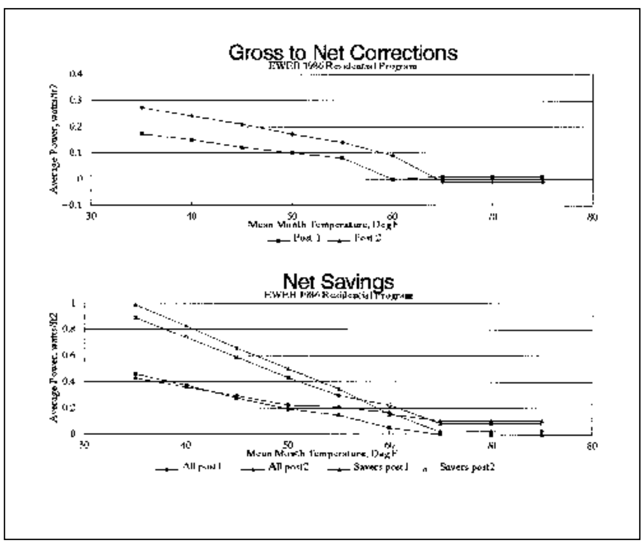


Figure 4. Gross to net corrections and net savings



ACKNOWLEDGMENTS

The authors wish to acknowledge and thank EWEB for their support and assistance with this project. The ''institutional memory'' of staff was most helpful. The EWEB program is an exemplary residential electrical energy conservation program. We also wish to acknowledge and thank Linda Wigington, Wigington & Associates, for her technical guidance and work with the participant group field visits. And, thanks to Delta-T, Inc. for their work with the ducted system homes and help with comparison group field visits.

Table 2. Net Levelized Cost of Savings EWEB 1986 program

Whole Sample, N = 20	97 mills/kWh
Sample size	26,378 square feet
Total cost	$53,838 1986$
Net total savings	42,675 kWh/yr first 3 years
Net total savings	47,156 kWh/yr years 4 to 6
Probable Savers, N = 9	41 mills/kWh
Sample size	10,712 square feet
Total cost	$19,258 1986$
Net total savings	42,734 kWh/yr first 3 years
Net total savings	38,163 kWh/yr years 4 to 6

ENDNOTES

1. The Bonneville Power Administration has run a Regional Weatherization Program since 1980, weatherizing approximately 280,000 single family residences with permanently installed electric space heating equipment.

2. Brown, Marilyn and Dennis White. December 1992. *Evaluation of Bonneville's 1988 and 1989 Residential Weatherization Program: A Northwest Study of Program Dynamics*. ORNL/CON 323. Oak Ridge, TN: Oak Ridge National Laboratory. In addition, yearly Bonneville evaluation studies were also reviewed, from *Evaluation of the BPA Residential Weatherization Pilot Program*, 1983, ORNL/CON-124, through ORNL/CON 323.

3. EWEB provided access to customer files which included audit data, work authorizations, inspection data, cost and buy back data. Consumption records were accessed. Staff members who had worked with the RWP since 1986 were also interviewed.

4. Called the ''Data Gathering Project,'' Bonneville gathered consumption, measure installation, and cost data from various utilities, located across the regions' climate zones, which were used for the impact evaluations conducted by Oak Ridge National Laboratories, such as those cited above. Selecting homes for this study which were included in previous ORNL studies provides a link to the ORNL studies and the opportunity to look closely as some of these homes.

5. The general population of participants had higher wood usage than was allowed in this sample. Wood use was

restricted in this sample to limit the ''noise'' in the billing data caused by auxiliary wood heat.

6. PRInceton Scorekeeping Method. The Center for Energy and Environmental Studies, Princeton University. Both the DOS and Advanced Version 1.0 were used.

7. Weather normalization is based on the least squares fit to billing and temperature data. A comparison between the manual fit used in this study and the PRISM method showed an almost identical fit when there were no outlying data points. Our inspections showed that most outliers were behaviorally caused. Removal of the outliers renders the fit more representative of the physical building, while including the outliers makes the fit more representative of buildings including occupant behavior.

8. See, for example, ORNL studies cited in endnote 2 above.

H. GIL PEACH & ASSOCIATES

MANAGEMENT/PLANNING/EVALUATION

16232 NW OAK HILLS DRIVE
BEAVERTON, OREGON 97006-5242 USA

Telephone: (503) 645-0716
Fax: (800) 204-3803
hgilpeach@scanamerica.net

Suggested Citation:

West, Anne Minor, Howard S. Reichmuth, Pamela Brandis, & H. Gil Peach
"Seven Years After: Impact Evaluation Results Employing Extensive Site Inspection Data and Associated Pre / Post Billing Analysis," Pp. 3.97-3.104 in the *Proceedings of the 1996 ACEEE Summer Study on Energy Efficiency in Buildings*, Panel 3, Residential Programs: Program Evaluation.