Supporting Statement
Weatherization Assistance Program Evaluation
OMB Control Number: (NEW)

B. Collections of Information Employing Statistical Methods

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used.

The Weatherization Assistance Program evaluation will require statistical sampling of weatherization agencies and, by subsampling, agency weatherization staff, weatherization and control home occupants, and weatherization and control home utility billing data. A listing of each weatherization agency along with its planned Fiscal Year (FY) 2008 dollar allocation and number of weatherized units was obtained from the WinSAGA (Systems Approach to Grants Administration for Windows) data base.[1] As of 2008 there were 904 agencies. Sampling will be stratified by state with probability proportional to agency FY 2008 dollar allocation. Table B.1 shows the FY 2008 WinSAGA total for each state.

In addition, 100% of State (and the District of Columbia) weatherization program managers or their delegates will be surveyed.[2, 3] States will be sampled completely as per OMB guidance: "…a census of state program directors may provide higher quality information with little cost difference from a sample survey of a slightly smaller number of states. In this case, there may also be concerns about missing practices of some states that were not included in the sample if a census were not conducted."[4]

With a few refinements and additions, the sampling will be done essentially to emulate the last Weatherization Assistance Program evaluation, which was conducted in 1993 using data from the 1989-1990 program year.[5] In that study, 400 agencies were sampled and 361 responded. Occupant level data were obtained by simple random subsampling of

---

[1]WinSAGA data kindly provided by Christine Askew, Office of the Weatherization and Intergovernmental Program, Energy Efficiency and Renewable Energy, U.S. DOE.

[2]DOE Weatherization Assistance Program management will also be sampled. That sampling is not considered as part of this Information Collection Request (ICR) because, as discussed in 5CFR1320.3, the information acquired is to be used specifically for the Weatherization Assistance Program evaluation and not for general statistical purposes.

[3]State weatherization program contact information may be found at http://www.eere.energy.gov/weatherization/state_contacts.html.

[4]"Questions and Answers When Designing Surveys for Information Collections," Office of Information and Regulatory Affairs, Office of Management and Budget, http://www.whitehouse.gov/omb/inforeg/pmc_survey_guidance_2006.pdf, January 2006.

[5]Brown, Marilyn A., Berry, Linda G., Balzer, Richard A., and Faby, Ellen, "National Impacts of the Weatherization Assistance Program in Single-Family Dwellings," ORNL/CON-326, Oak Ridge National Laboratory, Oak Ridge, Tennessee, May, 1993 (http://weatherization.ornl.gov/pdf/ORNL_CON-326.pdf).

one third of records for the weatherized units and one third of records for the control units for each responding agency. This rate applied to single-family detached dwellings and mobile homes. To properly determine weatherization savings in multi-family buildings, all units in each selected building must be analyzed. Therefore, multi-family buildings were selected at the rate of one third of buildings.

Table B.1. Weatherization Agencies
Source: FY 2008 WinSAGA Data

State	N	National Allocation ($)	PlannedNumber of Units
AK	5	1,542,283	456
AL	16	2,720,538	753
AR	15	1,764,892	533
AZ	11	1,382,049	770
CA	41	5,636,152	3,252
CO	9	11,787,683	3,275
CT	5	2,927,151	830
DC	5	574,163	175
DE	2	1,488,172	431
FL	31	1,817,557	602
GA	21	2,487,736	662
HI	4	187,272	115
IA	18	4,349,651	1,074
ID	7	1,847,343	432
IL	35	10,905,174	3,876
IN	24	6,360,587	1,776
KS	8	2,224,142	571
KY	23	3,326,960	924
LA	20	7,148,952	1,953
MA	12	6,052,272	2,623
MD	15	2,904,363	980
ME	10	2,860,300	824
MI	32	16,055,393	5,000
MN	33	9,719,154	2,916
MO	18	6,475,487	1,954
MS	9	1,403,639	414
MT	11	9,539,215	2,262
NC	33	3,785,388	994
ND	7	4,174,166	980
NE	9	2,372,781	601
NH	6	1,383,471	396
NJ	22	4,713,515	1,261
NM	4	2,262,283	777
NV	4	761,347	189
NY	66	58,719,241	12,800
OH	59	14,554,002	6,694
OK	19	2,963,385	820
OR	21	3,241,349	891
PA	42	13,655,395	3,798
RI	7	947,703	243
SC	9	1,600,719	443
SD	4	1,887,890	547
TN	18	8,285,736	2,984
TX	34	5,258,169	1,417
UT	8	2,091,958	554
VA	22	3,622,545	901
VT	5	1,126,565	797
WA	26	3,914,348	1,092
WI	22	7,040,088	2,341
WV	13	3,178,264	1,291
WY	4	1,526,687	380
ALL	904	278,555,274	82,624

Because the weatherization agencies hold the occupant and utility contact information, the subsampled populations can only be listed for those weatherization agencies to be sampled and only after agency sampling is performed. Therefore, no listing of subsampled populations is made here. However, Table B.2 shows the proposed sample sizes for the subsampled components. The basis for the proposed sample sizes is discussed in this section and Sections B.2.3.1-5. Note that some of the study components will overlap in that the same subsampled units may be used for more than one study. Thus the totals in Table B.2 do not necessarily add to actual totals needed for the study.

Table B.2. Weatherization Assistance Program Evaluation Proposed Sample Sizes for Sampled Study Components

Survey Component	Proposed Sample Size	Basis Discussion (in this document)
Agencies	400 agencies	B.1 (this section)
SF Fuel-oil-heated	128 occupants*	B.2.3.1
SF Propane-heated homes	128 occupants*	B.2.3.2
MF Fuel-oil-heated homes	20 building supervisors	B.2.3.3
Propane-heated mobile homes	128 occupants*	B.2.3.4
Air conditioned homes	264 occupants*	B.2.3.5

*Homes will be used for multiple component studies when possible without causing bias.

Agency nonreponse is biasing, but was only 10% ($100\times(1 - 361/400)$) in the 1993 evaluation and may be substantially less for the proposed evaluation because today's more efficient electronic record keeping makes complying with data requests easier than it was previously. In addition to agency nonresponse, a 50% nonresponse rate by utilities was encountered in requests for billing data.[6] Utility nonresponse is nonbiasing because it does not reflect on the weatherization agencies themselves. The issue of utility nonresponse is important, however, because it reduces the final acquired sample size.

For the present evaluation, a sample of 400 agencies stratified by state will be selected by probability proportional to size (PPS) sampling and 2008 funding allocation as the measure of size. (In addition, all 904 agencies will be asked to fill out short forms about activities in PY 2007 and 2008.) To assess the effects of nonresponse and the nature of the sampling in general, a pseudorandom sample was selected from the WinSAGA listing assuming 10% nonresponses.[7] The sample ultimately selected will be similar. Table B.3

[6]Ibid.

[7]Sampling done using the SAS Survey select procedure (2004 SAS/Stat 9.1 User's Guide, Cary, NC: SAS Institute, Inc.)

below shows totals for this sample. The stratified PPS sampling procedure is discussed in part B.2.1.

Table B.3 also shows the number of units actually sampled (16,230) in a one-third subsample of units, and, assuming (as in the 1993 evaluation) that 60% of units are gas or electric, the number of units (9,738) potentially available for the billing analyses. Utility non-response will depreciate this number. However, the final column in the table shows that if data are obtained for 49% of the units potentially available for billing analysis, then the number of units sampled will equal the number sampled in the 1993 evaluation. Because the 49% acquisition rate is slightly less than the response rate seen in the 1993 evaluation, the sample ultimately acquired for the present evaluation should be as large as or slightly larger than the previous sample.

Table B.3. Weatherization Assistance Program 2008 National Totals for a PPS Sample Stratified by State with PPS "Size" = 2008 Dollar Allocation and 10% Nonresponse*

National Agencies (2008 Listing)	Agencies Sampled	Agencies Responding	Percent of National Agencies Responding	2008 National Allocation	Allocation Represented by Responders	Percent Allocation Represented by Responders
904	400	361**	40%	$278,555,274	$175,315,062	63%

2008 National Units Planned	Units Represented by Responders	Percent of Units Represented by Responders	Units Subsampled (at 33%)	Gas/Elec Units Subsampled (60% Approx)	Capture Rate To Achieve 1993 Number Usable (4,796)
82,624	49,208	60%	16,403	9,842	49%

*Assumes 10% nonrespondents selected as every tenth sampled agency; not necessarily the actual nonrespondents

**Because of constraints on PPS sampling, the two largest agencies were sampled with certainty. With every tenth sampled agency taken as a nonresponder, 361 agencies were assumed to respond (as in the 1993 evaluation).

In addition to energy use and savings data, information about the weatherization process and program, also necessary for the evaluation, will be obtained from the 400 sampled agencies. Sampled agencies will be asked to provide agency staff contact information and functional classifications, weatherized and control building/unit occupant lists, and dwelling data. Weatherized and control unit dwelling data will be obtained using the one-third unit/building subsampling as discussed above for the 1993 evaluation.

2. Describe the procedures for the collection of information including:

B.2.1. Statistical methodology for stratification and sample selection

Agency sampling will be stratified by state with PPS sampling within states and with 2008 funding allocation as the PPS sampling size measure. Allocation of sample to individual states will also be in proportion to size. Table B.3 shows totals for a possible sample of agencies. Sampling weights will be computed when the sample is generated. Subsampling (of occupants, agency crew, etc.) will be by simple random sampling (SRS) in some cases stratified (e.g., by crew function, weatherized-vs-control subjects).

B.*2.2. Estimation procedure*

Most of the statistical analysis will be to compute summary means and frequencies. All such analyses will account for the stratified sampling design and sampling weights (e.g., with the SAS[8] Surveymeans or Surveyfreq procedures). However, billing data and other fuel consumption analyses require regression analysis to adjust for differences in weather across seasons and locations. These adjustments will be made using multiple approaches (as an internal check), for example the Princeton Scorekeeping Method[9] and as well as a straightforward linear model approach with billing days and heating and cooling degree days as independent variables.

B.2.3. Degree of accuracy and sample sizes needed for the purpose described in the justification

The proposed evaluation is designed to substantially emulate the last Weatherization Assistance Program evaluation, which was conducted in 1993.[10] Estimates and standard errors of the mean control-adjusted natural gas and electricity savings per weatherized unit per year, which are two of the primary endpoints in the 1993 evaluation, are listed in the following table:

Primary Heating Fuel	Average Savings per Weatherized Unit (1)	Standard Error of Average (2)	Relative Error: (2)/(1) as Percent
Natural gas	17.8 million Btu (MMBtu)/year	1.8 MMBtu/year	10%
Electricity	1,830 kilowatt-hours (kWh)/year	358 kWh/year	20%

[8]2004 SAS/Stat 9.1 User's Guide, Cary, NC: SAS Institute, Inc.

[9]Fels, M., K. Kissock, M. Marean, and C. Reynolds, "PRISM Advanced Version 1.0 User's Guide," Princeton University, Center for Energy and Environmental Studies, Princeton, NJ, 1995.

[10]Brown et al, op. cit.

The precision obtained in the 1993 evaluation was subsequently found to be adequate for that evaluation and is assumed adequate for the proposed evaluation as well.

However, the proposed evaluation will incorporate several refinements and additions. The 1993 evaluation was stratified by agency size and geographic region. The agency size strata were sampled at the same rate except for the largest-size stratum which was certainty sampled. Thus, with the exception of the very largest agencies, large agencies had no greater chance of selection than small ones. Yet, in general, the larger the agency the greater its contribution to total energy savings. In the proposed evaluation, PPS sampling with size measured as agency funding will provide a refinement of the 1993 stratification by agency size by allowing agency size to be continuously reflected in the sampling probabilities, with the effect that all agencies will be statistically represented, but larger agencies will be sampled preferentially.

The geographic stratification in the proposed evaluation will also be a refinement of the approach used in the 1993 study. The earlier study employed ten climate subregions, which were approximations of standard climate regions based on state boundaries. For the proposed study, in addition to representing all climate zones, it was considered politically advantageous to guarantee representation of all states. Therefore, stratification for the proposed study will be by state.

Although the sampling for the proposed study is a refinement of the 1993 sampling, the two study designs are substantially similar, and the 1993 study is by far the best available source of prior information for the proposed one. Response rates for the proposed evaluation are expected to be as good as, or better than, the rates seen in 1993. Therefore, for the purpose of sample size calculations, it is reasonable to regard the proposed evaluation as emulating the 1993 study. Thus, as in the 1993 evaluation and as discussed in part B.1 of this document, 400 agencies will be sampled in the proposed study.

This defines the sample size requirement for the primary sample of agencies. With the exception of the complete sampling of states, all other components of the evaluation will require subsampling through the primary agency sample. The following subsections of this section describe the sample size requirements for these various subsampled component studies.

B.2.3.1. Single-family fuel-oil-heated homes

For bulk-delivery fuels such as fuel oil and propane billing, delivery amount records are too discrete for energy analyses. In-home metering is required for such bulk fuel studies. In-home metering studies are expensive in comparison with billing-data analysis. However, because most weatherization measures are independent of fuel type (e.g., insulation, sealing), it is reasonable to assume that weatherization energy savings (though not necessarily dollar cost savings) do not vary much among fuel types. This assumption

is supported by the fuel-oil component of the 1993 evaluation.[11] However, because overall program savings estimates computed under this assumption will obviously depend on it, it is a good idea to test it for the major fuel alternatives to natural gas and electricity, namely fuel oil and propane.

The objective of the fuel oil (and propane) single-family study then will be to test the hypotheses that the weatherization energy savings for homes heated with these fuels are the same (on average) as the savings for natural gas heated homes. More formally, the null hypothesis H_0: "Mean savings per unit for fuel oil is same as for natural gas," will be tested against the alternative that the mean savings is different, with a probability of at least .90 of detecting a difference of 20% or more of the mean natural gas pre-weatherization normalized annual consumption (NAC).

The study of fuel-oil heated homes conducted as part of the 1993 evaluation will serve as a pilot study for reckoning sample sizes for the fuel oil as well as the propane single-family components of the proposed evaluation. Single-family fuel oil heated homes will be subsampled from lists provided by agencies that weatherize appreciable numbers of single-family fuel-oil-heated homes. The 1993 evaluation was conducted similarly by sampling agencies and subsampling weatherized and control units of the agencies. Because the hypothesis being tested was considered in the previous study, and because the proposed study has an additional component for propane (as well as refrigerators and air conditioners), the proposed fuel oil study will be a scaled-down version of the 1993 Fuel Oil evaluation. The scale factor is suggested as follows.

Let A_N and A_F denote estimates (to be computed once data are collected) of the average per-unit savings for natural gas and fuel oil respectively. Let SE_N and SE_F denote the standard errors of these estimates. As discussed above, SE_N is expected to be approximately the same for both the proposed study and the 1993 evaluation. Let $f = N_P/N_{90}$, where N_{90} is the number of units sampled in the 1993 fuel oil study (approximately 300, see below) and N_P is the number of units (to be determined) in the proposed study. Then the standard error of the difference $D = A_N - A_F$ for the proposed study is $SE_D = [(SE_N)^2 + (SE_F)^2/f\,]^{1/2}$.

Again let $Z_{.95} = 1.64$ denote the 95th percentile of the standard normal distribution, and let $T = D / SE_D$. An approximate level 0.1 test of H_0 can be conducted by rejecting H_0 for $| D / SE_D | > Z_{.95}$. For this test, $P(\text{ Reject }) = 1 - P(-Z_{.95}\, SE_D < D < +Z_{.95}\, SE_D)$. For a given true difference Δ, $P(\text{ Reject }) = 1 - P(-Z_{.95}\, SE_D - \Delta < D - \Delta < Z_{.95}\, SE_D - \Delta) = 1 - P(- Z_{.95} - \Delta / SE_D < (D - \Delta) / SE_D < Z_{.95} - \Delta / SE_D) = 1 - \Phi(Z_{.95} - \Delta / SE_D) - \Phi(- Z_{.95} - \Delta / SE_D)$, where Φ denotes the cumulative distribution function of the standard normal distribution.

From the 1993 evaluation, the average annual fuel savings (± standard error) are 22.4 (± 2.7) MMBtu. This was for 193 fuel-oil-heated weatherized units and 105 fuel-oil-heated

[11]Levins, William. P., and Ternes, Mark P. (1994). "Impacts of the Weatherization Assistance Program in Fuel-Oil Heated Houses," ORNL/CON-327, October 1994 (http://weatherization.ornl.gov/pdf/0RNL_CON_327.pdf).

control units (approximately 300 fuel-oil-heated units in all) selected from 41 agencies. The pre-weatherization NAC of natural gas is 137.4 MMBtu.[12] Taking Δ to be ten percent of the pre-weatherization NAC (estimated), Δ = 13.74 MMBtu. For this Δ, P(Reject) can be computed for various values of *f*. As *f* increases, so does P(Reject). A SAS program was written to calculate P(Reject) for each *f*. It turns out that P(Reject) = .90 for *f* = .39. This suggests that a survey 39% as big as the 1993 fuel oil study is needed for the proposed fuel oil study: about 75 weatherized and 41 control homes, 116 homes in all from about 16 agencies. Because the variability of savings (i.e., pre-post differences) of weatherized and control fuel-oil heated homes is about the same,[13] it is more efficient to take half weatherized and half control homes, that is, 58 of each rather than 75 and 41. This will be implemented by random sampling from single-family fuel-oil-heated homes identified in the dwelling information data provided by the sampled agencies that weatherized appreciable numbers of fuel-oil heated homes.

As a scaled version of the fuel-oil component of the 1993 evaluation, nonresponse is accounted for in the sample size calculation. However, prior experience[14] with in-home fuel-oil metering studies has shown that metering instruments fail or are damaged about ten percent of the time. This kind of nonresponse can be considered random and nonbiasing. The extent to which this censoring occurred in the 1993 study is unclear from the documentation. Therefore, to ensure an adequate sample size, increasing the sample size by 10% seems advisable. Thus 64 (= 58×1.1) weatherized and control homes will be sampled.

This study will be coordinated with the Air Conditioned Homes study (see Section B.2.3.5) to use common homes where possible without biasing the sampling design.

B.2.3.2. Single-family propane-heated homes

The hypothesis discussed in Section B.2.3.1 that weatherization savings are the same for fuel-oil and natural-gas heated homes is also reasonable for propane-heated homes. The propane component of the proposed evaluation is intended as a check on this hypothesis for propane. Because the underlying hypothesis is that the distributions of savings are the same for all three fuel types, it is reasonable to assume that the sample size for the proposed fuel oil study is also appropriate for the propane study. Therefore, as above for the fuel-oil component, 64 weatherized and 64 control propane-heated units will be sampled for the propane component, from agencies PPS-subsampled from agencies that weatherize appreciable numbers of single-family propane-heated homes. This will be implemented by random sampling from single-family propane-heated homes identified in the dwelling information data provided by the agencies.

This study will be coordinated with the Air Conditioned Homes study (see Section B.2.3.5) and the Direct Measurement of Selected Household Factors study (see Section B.2.3.6) to use common homes where possible without biasing the sampling design.

---

[12]Brown et al, op. cit.
[13]Levins and Ternes, op. cit., Table ES.1.
[14]Ibid.

B.2.3.3. Multi-family fuel-oil-heated homes

An approach similar to that taken for single-family fuel-oil-heated homes (Section B.2.3.1) can be taken for multi-family fuel-oil-heated homes. Preliminary data about multi-family homes are available from the 1993 evaluation. The objective of the multi-family home component of the proposed evaluation will be to test the hypothesis that the weatherization energy savings in fuel-oil and natural-gas heated multi-family homes are the same.

However, multi-family buildings pose a more complicated analysis problem than single-family homes, because (1) there are fewer multi-family buildings than single-family homes, (2) multi-family buildings vary substantially in numbers of individual home units, and (3) to properly understand the effect of weatherization on multi-family buildings, all of the dwelling units in each multi-family building sampled must be analyzed collectively. Because multi-family buildings vary substantially in numbers of individual home units (for example, the thirteen natural gas heated multi-family homes examined in the 1993 study varied from 6 to 80 in numbers of units), consumption and savings per unit will be considered as the metric of interest, rather than savings *per se*.

For single-family fuel-oil heated homes, as discussed above, control-adjusted fuel-oil savings estimated by metering will be compared to control-adjusted natural gas savings (estimated from billing data). Although from the accuracy alone it may be preferable to use control-adjusted savings (e.g., because of differences in natural gas and fuel-oil prices), savings for control homes tend to be small, and *differences* in control savings therefore tend to cancel out in control-adjusted savings calculations. Therefore, because of the complexity of, and smaller sample sizes used for, the multi-family fuel-oil component, no control adjustment will be made. That is, in estimating the difference between fuel-oil and natural-gas heated home weatherization savings, only weatherized homes will be compared.

The 1993 evaluation serves as a pilot study. Thirteen natural-gas heated multi-family buildings sampled in the 1993 study had a mean savings of 18.1 ± 4.4 (standard error) MMBtu per unit. The average pre-weatherization consumption per unit for these homes was 82.5 MMBtu. Assuming that the standard error for fuel-oil heated multi-family buildings is the same, the sample size needed to detect with 90% confidence a difference between the natural gas and fuel oil populations that is 20% of the natural gas pre-weatherization consumption can be calculated. A SAS program similar to the one written for the single-family fuel-oil-heated homes calculation shows that this factor is 1.5, which indicates a sample size of 20 ($= 1.5 \times 13$) for the proposed study. Thus 20 fuel oil multi-family fuel-oil heated buildings should be sampled. However, prior experience with multi-family building studies has shown that data obtained for them will be inadequate about 20% of the time, and the extent to which this occurred in the 1993 study is unclear from the documentation. Therefore, to ensure an adequate sample size, increasing the sample size by 20% seems advisable. Thus 24 ($= 20 \times 1.2$) multi-family buildings will be sampled. (Note that this kind of nonresponse can be considered random and nonbiasing.)

The sampling will be implemented by random sampling from multi-family fuel-oil-heated homes identified in the dwelling information data from agencies that weatherize multi-family fuel-oil-heated homes. Agency sampling weights will be accounted for in the data analysis.

B.2.3.4. Propane-heated mobile homes

Although weatherization savings in mobile homes differ from savings for single-family site-built homes, it is reasonable to regard the savings for propane-heated mobile homes as about the same on average as the savings for natural-gas heated mobile homes. Because of billing data analysis, the natural gas savings is much easier to estimate. Further, the sample size needed to assess the difference (if any) between fuel types in mobile-home energy savings should be about the same as the sample size needed to assess the difference between fuel types in conventional homes. This suggests a sample size of 64 weatherized and 64 control units, from agencies that weatherize appreciable numbers of propane-heated mobile homes, as suggested for the single-family fuel oil study. This will be implemented by random sampling from propane-heated mobile homes identified in the dwelling information data provided by the agencies.

This study will be coordinated with the Air Conditioned Homes study to use common homes where possible without biasing the sampling design.

B.2.3.5. Air conditioned homes

Because of the wide variability in estimated weatherization savings for warm-climate states, it has been hypothesized that weatherization in warm-climate states does not achieve any air conditioning energy savings at all. In one air conditioner (AC) study, for a sample of 22 weatherized homes, the mean AC energy savings was –31 kWh, with a standard error of 167.2 kWh.[15] The mean AC savings for a sample of 19 control homes was 106.7 with a standard error of 112.1. These results are consistent with the hypothesis of no AC savings. This study was done in Oklahoma, but results were similar in a study of AC savings in North Carolina.[16]

The object of the proposed AC study is to test the hypothesis that mean AC savings in warm-climate states are zero (H_0: Mean AC Energy Savings = 0) against the alternative that the mean savings are positive, with a probability of at least .90 of detecting a savings of ten percent of the pre-weatherization AC consumption. The Oklahoma study mean pre-weatherization AC consumption combined estimate from both weatherized and control groups is 1,652.4 kWh, ten percent of which is 165.2 kWh.

[15]Ternes, Mark P., and Levins, William P. (1992), "The Oklahoma Field Test: Air-Conditioning Electricity Savings from Standard Energy Conservation Measures, Radiant Barriers, and High-Efficiency Window Air Conditioners," Oak Ridge National Laboratory, ORNL/CON-317, August 1992 (http://weatherization.ornl.gov/pdf/ORNL_CON_317.pdf).

[16]Sharp, T. (1994), "The North Carolina Field Test: Field Performance of the Preliminary Version of an Advanced Weatherization Audit for the Department of Energy's Weatherization Assistance Program.," ORNL/CON-362, June 1994 (http://weatherization.ornl.gov/pdf/ORNL_CON-362.pdf).

The proposed study will be conducted through weatherization agencies, the primary sampling units. As an approximation in reckoning sample sizes, we ignore the agency sampling weights (but they will be accounted for in the analysis). Assuming a level 0.1 one-sided hypothesis test, and using the Oklahoma study for preliminary estimates of the standard error and pre-weatherization AC savings, the sample size necessary for detecting a weatherization effect of 165.2 kWh or more can be estimated as follows.

From the weatherized and control group sample sizes (22 and 19) and standard errors (167.2 kWh and 112.1 kWh), it can be shown (using an F-test) that the sample variances are not significantly different. Therefore, a pooled standard deviation will be used, and the same sample size will be assumed for both weatherized and control groups. The pooled standard deviation estimate is $[(22\times21\times(167.2)^2 + 19\times18\times(112.1)^2)/(21 + 18)]^{1/2}$ = 664.4. The standard error for weatherized-control-group difference of mean AC savings can therefore be estimated be as $664.4/N^{1/2}$, where N weatherized and N control units are to be sampled (2N units in all).

The usual one-sided normal-theory test at the 0.1 level rejects the null hypothesis when the difference of means divided by the standard error (SE) of the difference exceeds the .90 normal quantile $Z_{.90} = 1.28$. For a true mean difference of 165.2 (i.e., ten percent of pre-weatherization NAC). P(Reject) =P[difference/SE > 1.28] = P[(difference – 165.2)/SE + 165.2/SE > 1.28] ≈ P[Z > 1.28 – 165.2/(664.4/$N^{1/2}$)] = 1 – P[Z <= 1.28 – (165.2/664.4)$N)^{1/2}$]. This implies N = 106. That, is 106 weatherized and 106 control homes will be needed for the air-conditioned study. The sampling will be implemented by random sampling from air conditioned homes identified in the dwelling information data provided by a PPS-subsample of agencies from warm-climate states.

However, prior experience[17] with in-home AC metering studies has shown that AC metering instruments may fail or be damaged up to twenty five percent of the time. (Note that this kind of nonresponse can be considered random and nonbiasing.) To ensure an adequate sample size, increasing the sample size by 25% seems advisable. Thus 132 (106×1.25) homes will be sampled in each of the treatment and control groups (264 homes total).

This study will be coordinated with the other occupant sampling studies (i.e., of Single-Family Fuel-Oil/Propane Heated Homes (Sections B.2.3.1-2), and the Propane-Heated Mobile Homes study (Section B.2.3.3) to use common homes where possible without biasing the sampling design.

B.2.4. Unusual problems requiring specialized sampling procedures

None.

B.2.5. Use of periodic (less frequent than annual) data collection cycles to reduce burden

[17]Ternes and Levins, op. cit.

The Weatherization Assistance Program evaluation is conducted occasionally, not annually. The last evaluation was in 1993, using data from the 1989-1990 program year.

3. Describe methods to maximize response rates and to deal with issues of non-response.

B.3.1. Maximizing response rates

Contacts with states, local agencies, and utilities will be established to help promote the data collection process. Assistance will also be sought from professional organizations and, in the case of utilities, from regulatory commissions. Data requests will be made to minimize the demand on respondents. For example, requests to the same party for multiple data installments will be coordinated to minimize the workload. Web survey forms will be made available for ease of use. Electronic data delivery will be encouraged, but data will be accepted in any standard format.

B.3.2. Methods for dealing with non-response

Previous studies have shown that nonresponse by utilities in requests for billing data is likely to be substantial. However, utility nonresponse does not reflect positively or negatively on weatherization agency performance. Therefore, utility nonresponse will be treated as nonbiasing in the data analysis. (Of course utility nonresponse is accounted for in the sample size requirements.)

Previous studies of weatherization agencies have shown that nonresponse by agencies will be minimal—around 10% or so for the 1993 evaluation. Because agencies (as well as utilities and other businesses) are today much more likely to store data in electronic formats than they were in 1993, we expect to incur even lower nonresponse rates in the proposed evaluation. Therefore, nonresponse by agencies will be tracked but not adjusted for (other than in sample sizes) in the data analysis.

Attrition of occupants due to moving is another form of nonresponse that is also largely independent of weatherization agency performance. It is accounted for in the sample size (Section B.2.3.1) but will not be adjusted for in the data analysis. Other forms of nonresponse are expected to be minor and are discussed above in sections B.2.3.1-5 on individual sample size requirements.

4. Describe any tests of procedures or methods to be undertaken.

Seven information collection instruments have been pretested on small samples (fewer than 10 from each sampled population) and in several cases revised on the basis of pretest responses. The following list shows the populations and associated pretest sample sizes:

- **Agencies**: three to five agency administrative personnel from different agencies; six questionnaires (DF10: All Agencies Information Data Form; S2: All Agencies

Program Information Survey; S2: Subset of Agencies Detailed Program Information Survey; DF2: Housing Unit Information Data Form; DF3: Building Unit Information Data Form)

- **States**: three to five state administrative personnel from different states; two questionnaires (DF1: All States Agencies Information Data Form and S1: All States Program Information Survey)

5. <u>Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s) or other person(s) who will actually collect and/or analyze the information for the agency.</u>

Richard L. Schmoyer, Ph.D. (Statistics, 1980), of Oak Ridge National Laboratory developed the statistical components of the evaluation plan. He can be reached at 865-946-1255; schmoyerrljr@ornl.gov. ORNL will provide oversight to the evaluation contractor that administers the survey and performs the data analysis.

Appendix to Supporting Statement Part B. Collections of Information Employing Statistical Methods

- *Four components deal with response rates from utilities: 1) expected utility response rates; 2) experience with EIA; 3) non-biasing of utility non-response; and 4) a non-response bias and analysis plan.*

1) Utility non-response rates – Part B of the Supporting Statement that ORNL submitted to OMB as part of this ICR was written in late 2006/early 2007. In that Statement, it is stated that a response rate in the range of 50% from utilities was expected. This response rate was derived, in part, from a conservative interpretation of the experience with utility non-response from the previous national evaluation of WAP that was conducted about two decades ago.[18] A re-review of the past study combined with recent experience in contacting utilities for state-level evaluations and experience with EIA data collections suggests that 50% utility non-response rate may be an overly conservative assumption.

As part of the previous national evaluation of WAP, the evaluation team created a list of approximately 1500 utilities to contact, based on an initial compilation of information provided by the local weatherization agencies. Careful review of this information revealed that many utilities were double-counted in this list and other problems were discovered. Eventually, this list of 1500 utilities was reduced to 926, which were contacted for billing histories. Of this list, 689 responded, for a response rate of 74%. The documentation of the previous project is vague on how the list was paired down, so at worst the response rate was 46% and at best it was 74%. It is our belief now that the actual response rate was between 60-65%.

Admittedly, it is somewhat problematic to base utility non-response estimates on such dated information. The evaluation team that ORNL has assembled to conduct this retrospective evaluation of WAP has extensive recent experience in collecting billing histories from utilities. Table 1 below is a summary of the team's experience in collecting information from natural gas and electric utilities for four state-level evaluation projects.

The recent evaluations conducted by the ORNL team in these four states contacted 12 different utilities; 11 responded for a response rate of 92%. Without a doubt, this high level of response was due, in part, to states' interest in the evaluations and states' work to obtain utility cooperation for the studies. The high level of response was also due, in part, to working relationships that members of the ORNL developed with state agencies and utilities in recent years. These relationships will be valuable in helping to increase utility response rates for this evaluation.

The utilities contacted by these studies were not chosen at random. They were chosen because they serve the most customers. Thus, a second important insight to be gained from Table 1 is that a small number of (investor-owned) utilities serve the majority of natural gas and electricity clients in the United States. For example, only three natural gas

[18] Brown et al, 1993, see http://weatherization.ornl.gov/pdf/ORNL_CON-326.pdf

utilities in Iowa and Colorado serve 96.5% and 93% of natural gas clients in those states, respectively. The situation is approximately the same for the electric utilities in those states. In Wisconsin, a fifth state where the ORNL team has worked extensively, five natural gas and five electricity utilities serve 99% and 90% of clients in state, respectively.

It has been the experience of the team that large utilities are more responsive than small utilities to requests for billing histories. This is because the large utilities have excellent electronic billing records database systems and the staff to respond to these types of requests. Another reason is that the large utilities are frequently asked to work with state agencies and public utility commissions on research and policy analysis projects that require utility data as inputs. Thus, it is possible that a utility response rate as low as 50% could yield billing data for up-wards of 80-90% of the weatherized homes that heat with natural gas and electricity in our sample if the non-response rate for large utilities is low.

Table 1. Recent State-Level Evaluation Experience with Utility Response

	Iowa (2008)	Vermont (1998/1999)	Colorado (Date 2002-2004)	Ohio (Date 1994)
Number of Gas Utilities in State	26	1	4-16	>16
Number Contacted	3	1	3	5
Number Responding	3	1	3	4
% of clients served by those utilities	96.5	100	93	87
Number electric utilities serving WAP clients	14	*	50	>10
Number Contacted	2	*	2	5
Number Responding	2	*	2	4
% of clients served by those utilities	84	*	66	67

* Efficiency Vermont provided all needed electricity bills. No utilities were contacted.

2) Experience with EIA – One member of ORNL's evaluation team, APPRISE, Inc., has over twenty years of experience with EIA Energy Supplier surveys. EIA's approach to utilities, which is essentially adopted by our project and is described in the DF5 cover letter, has yielded very high response rates for such surveys. For the 1984, 1987, 1990, 1993, and 1997 RECS Energy Supplier Surveys, these procedures resulted in a 100% response rate for electric and gas utility companies and over a 95% response rate for delivered fuel vendors (fuel oil and LP companies). For the 2001 RECS Energy Supplier Survey, the company response rate was 100% for all types of companies. It should be noted that response to the RECS survey is mandatory. The RECS survey legislation allows EIA to levy fines (of $3000 at most) on energy suppliers that do not respond. Without asking energy suppliers it is uncertain how much the threat of fines influenced their decisions to participate, however it is the subjective opinion of those on the ORNL team that have worked closely with EIA over the years that EIA rarely required reference

to that authority to achieve high response rates. Thus, recent EIA experience with utility response rates suggests that our utility response rate will be high.

It appears that the next RECS data collection effort may overlap with our utility data collection efforts. The ORNL team has already started working with the RECS manager – Chip Berry – to discuss ways to communicate effectively with electric and gas utilities so that those organizations clearly understand the reason for the two different data collection agencies.[19] We will continue to share information with EIA to take advantage of any possible synergies between the two studies.

3) Non-biasing of Utility Non-response – Because utility behavior has no bearing on energy savings in homes, utility non-response is unlikely to inject bias into national energy savings attributable to WAP. One major variation in energy savings is local weatherization agency performance, specifically the performance of their in-house and/or contractor crews in installing weatherization measures. Agency non-response could inject bias into national energy savings estimates but based on previous experience we expect agency response rates to be extremely high. In any case, there is no interaction between agencies and their crews and local utilities in the weatherization process. Utilities have no influence at all over crew performance. Additionally, any utility non-response will have roots in internal issues, such as lack of available staff, data access or internal policies on data sharing, and in no way be tied to agency crew performance issues. Thus, while we believe our previous estimate of 50% utility non-response is probably too conservative by a large margin, even such a response rate would not inject bias into national energy savings estimates.

It should be noted that the previous national evaluation of WAP did address non-response bias (see Appendix E, pages E.12-13 of the 1993 document; Note that non-response is usually referred to as attrition in weatherization studies because of the ultimate focus on energy bills collected.) Comparisons were made of dwellings for which billing records were obtained or not obtained. The 1993 document (e.g., Table E-3) focuses on characteristics of the dwellings themselves, including type (detached, mobile, multi-family, etc.) and heating system type (central, non-central), and on measures (insulation, air-leakage control, water heater measures, etc.). Although characteristics of utilities (e.g., numbers of customers, rural/urban status) serving the dwellings may have been considered in the 1993 analysis, they are not discussed in the 1993 report. The reason for this is that although utility non-response might be coincidentally associated with program performance, causal association is unlikely, because utility decisions to respond or not are based on issues such as work burden or customer privacy, not agency performance. Therefore, the previous analysis focused on characteristics other than utility characteristics that were considered of greater concern as potential non-response bias issues.

4) Non-response bias and analysis plan – Nevertheless, in the proposed analysis, utility characteristics (e.g., numbers of customers, rural/urban status) will be considered in the

[19] Our project has adopted a distinct logo, a derivative of the nationally known weatherization logo, to help 'brand' our project communications with utilities and others.

suite of characteristics analyzed for association with non-response. Characteristics of agencies, residences, or utilities found or presumed to be related to performance, that are also found to be associated with non-response, will be considered as a basis for sampling weight adjustments. (Sampling weights in the 1993 analysis are discussed in Appendix E, pages E.1-11).

It is difficult to draw the line between utilities that might be slow responders and those that could be considered non-responders that through intensive efforts by the project, eventually responded. The project has in place a non-response follow-up plan that is essentially continuous in nature (see DF5 cover letter). However, the data collection time-line that was submitted to OMB at the same time as DF5 indicates that 'normal' non-response follow-up will take place over a four month period after initial utility contact and that 'intensive' non-response follow-up will continue for another four months. Utilities that have not responded after eight months will be considered non-responders. Thus, for the purposes of utility non-response analysis of potential bias, energy savings in weatherized homes served by utilities that responded within the first four months will be compared to energy savings in weatherized homes that responded during the second four months. If the overwhelming majority of utilities and the overwhelming number of billing histories are collected during the first four months, this analysis will not be conducted.

It should be noted that during the first four months, all non-responding utilities will receive that same amount of non-response follow-up attention. During the second four months, project resources will focus on the largest utilities first, and then on smaller utilities. Since our previous response to OMB on this matter, ORNL has received an assurance from DOE staff that requests for help in dealing with utility non-response will be made to the highest levels of DOE. The Secretary of Energy will intervene in situations involving the largest utility non-responders.

- *Additional explanation on design effects given that most sample designs involve clustering (i.e., the potential for increased variance in these substudies could create precision issues)?*

  The proposed evaluation was designed to be substantially equivalent to the previous evaluation (http://weatherization.ornl.gov/pdf/ORNL_CON-326.pdf) because the previous evaluation resulted in estimates and estimation precision that were ultimately considered both necessary and sufficient for evaluating the program. Both evaluations entail or will entail sampling 400 WAP agencies (clusters), with weatherized and control homes subsampled within agencies. Both evaluations were or will be analyzed using statistical methods that account for probability sampling and the stratified and clustered nature of the survey design. Agency nonresponse was minimal in the previous study, and the same is expected for the proposed study. Utility nonresponse, which was and likely will be more substantial, however, complicates the ultimate estimation precision.

Design effects, which relate sampling precision of complex stratified and clustered

surveys to the precision of resource-equivalent simple random samples, are used as approximations in survey analysis and sample size calculations. Because the proposed evaluation relates directly to the previous study, however, design effects were not used or needed in reckoning sample sizes for the proposed evaluation. Design effects were not computed or needed in the previous analysis, though they will be computed (incidentally with the software) in the analysis of the proposed study.

- *Analysis plan and assumptions in support of using of common homes where possible without biasing sample design.*

The WAP evaluation plan has specifications for six studies that involve direct measurement of conditions in client homes:

1) Fuel Oil Single Family Sub-Meter Study
2) Propane Single Family Sub-Meter Study
3) Propane Mobile Home Sub-Meter Study
4) Air Conditioning Study
5) Indoor Air Quality Study
6) Air Sealing, Duct Sealing, Furnace Study

The sample sizes and designs for each of these studies are discussed in Part B of the Supporting Statement.

The evaluation contractor has pointed out that the burden on survey respondents and the resources needed to implement the studies can be reduced by overlapping the samples from some of these studies. The following approach will be used to accomplish this while making sure that the overlap procedures do not compromise the statistical validity of the individual studies.

i) The samples for the three sub-meter studies (studies 1-3), which are statistically independent and from distinct target populations, will be selected first.

ii) The Air Conditioning (AC) study (study 4) will be conducted in the two hot regions with 33 of the 400 sample agencies. The AC study does not overlap with the fuel oil study (study 1), since that study is restricted to the nine Northeast States. It partially overlaps with the two propane studies (studies 2 and 3), since they are being conducted nationwide. The AC sample will be selected by first selecting a "pre-sample" for the AC study, equivalent to the sample that would have been selected if the AC study were being conducted alone. However, some members of the pre-sample are propane-heated and would qualify for study 2 (or study 3). Because all sampling here is random, they are statistically exchangeable with members of the study 2 (or study 3) samples that have AC and would qualify for study 4. That is, they can be randomly substituted for members of study 2 (or study 3) without compromising the validity of either study. Conversely, members from the sample for study 2 (or study 3) can be substituted for members of the AC study without compromising either study's validity. These substitutions will be made

to the extent possible, subject to the number of sampled members in the intersection of the target populations for both studies and on whether any of the studies are already underway.

iii) The Indoor Air Quality (IAQ) study (study 5) will be conducted nationwide. It overlaps geographically and conceptually with all three sub-meter studies as well as the AC study. Using the exchangeability-of-the-intersection approach discussed above, substitutions of members from the IAQ study for members of one of the other studies, or vice versa, will be made to the extent possible.

iv) The primary analysis in the Air Sealing, Duct Sealing, Furnace Work study (study 6) involves comparison of energy savings to measured performance from agency records. A secondary analysis involves an exploratory assessment of the data collected from homes in which other metering is being done. Since this study is designed to have a convenience sample (i.e., overlaps with other in-home visits), it will not be used to make formal statistical inferences about the performance of these procedures.