Evaluating the Impacts of Northwest Commercial New Construction Programs

Frederick D. Sebold, Regional Economic Research, Inc.
Boqing Wang, State of Washington

This paper describes an evaluation of a set of commercial new construction programs operated by five Northwest utilities. The programs in question are designed to influence construction decisions through the extension of engineering design assistance and/or financial incentives. They apply to a wide range of demand-side management (DSM) measures, although savings are dominated by lighting and HVAC measures. Both major remodeling jobs and build-outs, as well as new construction, are eligible for program services. All except one of these programs offer financial incentives for DSM activities.

The project had two major objectives. The first was to develop a rigorous method for evaluating the net impacts of new construction programs on participants' energy use patterns. The second was to demonstrate the use of this methodology by applying it to the five programs in question. Commercial new construction programs provide a particularly strong evaluation challenge because of the wide variety of measures they encompass, the diversity of commercial customers, variable occupancy during start up, and the absence of pre-adoption energy bills for statistical analysis. This paper discusses issues that must be confronted in the evaluation of the impacts of these programs, as well as the means by which these issues were addressed in this evaluation.

INTRODUCTION

This paper describes an evaluation of a set of commercial new construction programs operated by five Northwest utilities: Seattle City Light, Tacoma Public Utilities, Puget Power, Bonneville Power Administration, and Idaho Power. The programs in question are designed to influence construction decisions through the extension of engineering design assistance and/or financial incentives. They apply to a wide range of demand-side management (DSM) measures, although savings are dominated by lighting and HVAC measures. Both major remodeling jobs and build-outs, as well as new construction, are eligible for program services. All programs other than Idaho Power's offer financial incentives for DSM activities.

The evaluation was conducted under an Electric Power Research Institute tailored collaborative. The project had two major objectives. The first was to develop a rigorous method for evaluating the net impacts of new construction programs on participants' energy use patterns. The second was to demonstrate the use of this methodology by applying it to the five programs in question. Commercial new construction programs provide a particularly strong evaluation challenge because of the wide variety of measures they encompass, the diversity of commercial customers, variable occupancy during start up, and the absence of pre-adoption energy bills for statistical analysis. This paper discusses issues that must be confronted in the evaluation of the impacts of these programs, as well as the means by which these issues were addressed in this evaluation.

The paper is organized as follows. First, we discuss several general evaluation issues associated with the evaluation of new construction programs. Second, we outline the methodology developed in the course of the project and describe key evaluation results. Third, we summarize our findings and offer conclusions.

IMPACT EVALUATION ISSUES

Four interrelated impact evaluation issues confronted the evaluation team:

- **Defining Gross Savings.** New construction programs are multi-dimensional, covering multiple end uses and a variety of DSM equipment and measures that impact each use. Defining gross savings for participants and nonparticipants requires reference points. (Note that the eventual analysis of net savings requires information about nonparticipants' gross savings as well as participants' gross savings.) These gross savings are typically measured relative to common practice or the level of efficiency that would prevail from strict compliance with standards. However, this does not mean that common practice or standards comprise the overall baseline for the net impact evaluation; they merely comprise convenient intermediate references for the gross savings analysis.

- **Estimating Realized Gross Savings.** There are two traditional means of estimating gross savings for new construction programs: engineering analysis (perhaps calibrated to bills at the site level) and mixed engineering-statistical analysis. This latter approach, which offers insights relating to the general relationship between engineering estimates and realized savings, was chosen for this evaluation. This method essentially relates differences in consumption across buildings to different stocks of DSM measures (different levels of energy efficiency) in these buildings. Clearly, this is a difficult impact estimation problem, in that we must control for a wide range of factors affecting differences in energy use levels across buildings. Unfortunately, small sample sizes and the heterogeneity of commercial buildings can seriously handicap efforts to statistically control for other factors and isolate program impacts. This expanded problem of statistical control calls for a very highly structured estimation approach. Estimating impacts is even further complicated by the inherent variability of new building loads. In their first few years of operation, new buildings may undergo a variety of changes, including dramatic changes in occupancy, HVAC system calibration and commissioning, changes in operating modes, and additions to equipment stocks. Ideally, statistical analyses should recognize these sources of load variability over time.

- **Defining the Baseline for Net Savings.** Defining the baseline against which program impacts are measured is conceptually straightforward but difficult in practice. The true baseline is what participants would have done in the absence of the program, but this is not directly observable. As a result, evaluators sometimes use nonparticipant behavior as a proxy. However, the use of nonparticipants as a comparison group can result in significant bias in the estimation of net program effects.

- **Estimating Net Savings.** Some means of mitigating self-selection bias, as well as controlling for other differences between participants and nonparticipants, must be developed if net program savings are to be estimated. This entails the specification of a model of behavior covering both adoption decisions and participation decisions, and the derivation of a set of net-to-gross ratios for the affected end uses.

IMPACT EVALUATION METHODOLOGY

The methodology used to estimate the impacts of the Northwest New Construction Programs was comprised of the following elements: collection of an extensive amount of data on participants and nonparticipants, development of engineering estimates of gross savings associated with DSM measure adoptions, the development and application of a means of statistically calibrating these engineering estimates to billing records on the subject sites, development of a set of measures of realized energy efficiency for all participants and nonparticipants, application of econometric models of participation and efficiency choices to the estimation of net-to-gross ratios for each end use, and integration of the results of these analyses into a set of estimates of program impacts. These steps of the analysis are described briefly below.

Data Collection

Data collected for the impact evaluation included: on-site survey data for participants and nonparticipants, including data on building type, floor space by major space type, equipment inventories, shell characteristics, operating schedules, utility meter numbers, changes in space occupied or major equipment over time, and occupancy information; program tracking files, which contained information on measure costs and incentives, as well as ex ante estimates of savings; weather data for stations close to the surveyed sites; and monthly billing records.

Engineering Analysis

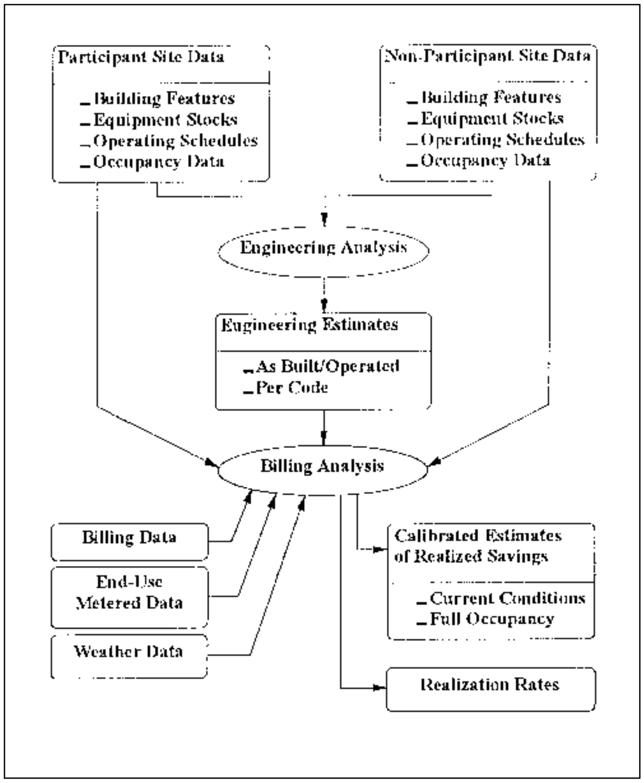
The engineering analysis was conducted by Architectural Energy Corporation (AEC), and yielded estimates of monthly energy consumption by end use under two scenarios: the *as-built* scenario, defined with respect to the actual equipment stocks and operating patterns found at the site; and a *reference* scenario, defined in terms of reference equipment and shell efficiencies (Jacobs & Roberts 1995). These reference values were defined in terms of either building codes or standard practices in the marketplace.

Statistical Calibration of Engineering Estimates of Gross Impacts

Generally, billing analysis is used to develop estimates of gross program savings on the basis of observed differences in energy usage associated with different levels of energy efficiency. As noted earlier evaluations, the analysis of new construction program impacts is plagued by a variety of practical problems, including the need to control for a wide range of other factors that differ across buildings, the need to accommodate small sample sizes, and the need to recognize factors causing variations in loads over time. The statistical approach used in this evaluation was a technique that can be called the realization rate approach.

The general logic of the realization rate approach (as applied to new construction programs) is illustrated in Figure 1.

Figure 1. *Realization Rate Framework*



The first step of the analysis entails the development of engineering estimates of end-use consumption levels. These estimates are based on information on building features, equipment stocks, operating schedules, and occupancy data. As shown, two types of engineering estimates are constructed for each site: estimates of end-use consumption under the reference assumption (minimal compliance with building standards or adoption of standard practice), and estimates of usage under the as-built scenario. Comparison of these two estimates yields the savings expected from departures from the reference options. The statistical (billing analysis) component of the realization rate model incorporates these engineering priors as well as information on site characteristics, weather conditions, and occupancy characteristics that might affect the realization of the engineering estimates of baseline usage and DSM-related savings. The model produces adjustment coefficients (or adjustment functions) that translate these engineering estimates into estimates consistent with observed energy usage and savings. These coefficients are called realization rates. As explained below, the realization rates on savings reflect the proportion of engineering-based savings estimates realized in the form of reduced site usage.

Model Specification. To derive the realization rate model, we begin with the standard statistically adjusted engineering (SAE) specification:

$$E_{bt} \;=\; \sum_e \alpha_e EEACTUAL_{bet} + \epsilon_{bt}$$

where E_{bt} is the actual energy usage at site b in billing period t, $EEACTUAL_{bet}$ is an engineering estimate of end-use consumption in the as-built scenario, and ϵ_{bt} is a random error associated with the building in period t. The presence of the adjustment coefficient α_e reflects the possibility of general engineering bias. The model can be expanded by decomposing the engineering estimates into two elements:

$$E_{bt} \;=\; \sum_e \alpha_e [EEBASE_{bet} - (EEBASE_{bet} - EEACTUAL_{bet})] + \epsilon_{bt}$$

where $EEBASE_{bet}$ represents an engineering estimate of usage under the reference assumptions with respect to the presence of energy conservation measures. This specification simply splits the engineering estimate into a baseline estimate and an estimate of the savings associated with the energy conservation beyond reference levels. Once the model is put into this form, possible modifications are apparent. First, the basic adjustment coefficient on the estimated energy savings can be allowed to be different from the adjustment coefficient of the baseline engineering estimate. Second, these adjustment coefficients can be permitted to vary across sites as conditions vary. One possible version of the revised model is as follows:

$$E_{bt} \;=\; \sum_e \alpha_e(X_{bt}) [EEBASE_{bet} - \beta_e(EEBASE_{bet} - EEACTUAL_{bet})] + \epsilon_{bt}$$

where β_e is an adjustment coefficient encompassing two phenomena: (a) the bias in engineering savings estimates *relative* to the bias in the reference energy usage estimates, and (b) the presence of behavioral rebound. Note also that the overall adjustment coefficient ($\alpha_e(X_{bt})$) is assumed to be a function of relevant factors. These factors could include site characteristics, weather, or other variables thought to affect the overall accuracy of baseline engineering calculations. In this application, the following site features were used in the estimation of the realization rate model:

- site square footage (used to normalize the model),
- building category binary variables (used to test for differences in realization rates across building categories), and
- occupancy rates (used to account for variations in occupancy).

Use of the Model to Infer Realization Rates. Once the parameters of this model are estimated, the end-use specific realized savings associated with differences between baseline efficiency levels and the as-built levels of efficiency in a building covered by the analysis would be:

$$\text{REALIZED SAVINGS}_{bet} = \hat{\alpha}_e(X_{bt})\hat{\beta}_e(\text{EEBASE}_{bet} - \text{EEACTUAL}_{bet})$$

where $\hat{\alpha}_e$ and $\hat{\beta}_e$ refer to estimated values of the parameters in question. Note that the associated realization rate for the end use and building is $\hat{\alpha}_e(X_{bt})\hat{\beta}$e.

There are several points to note about this approach:

- It makes full use of engineering estimates under baseline and high-efficiency scenarios. By doing so, it allows for at least some level of rebound.

- It can be used to account for changes in realized savings over time, and can generate estimates of steady-state (full-occupancy) savings.

- It provides a convenient means of adjusting engineering savings for errors associated with weather conditions for weather-sensitive end uses.

- Realization rates derived for a representative sample of participants are applicable to other participants subjected to the same engineering algorithms and assumptions.

Realization Rate Estimation Results. Estimated coefficients of the realization rate model are available upon request. Table 1 presents the estimated realization rates for heating, interior lighting, exterior lighting and refrigeration, as well as an overall weighted average rate. As shown, the realization rates for interior lighting, space heating, exterior lighting and refrigeration are above 1.0, suggesting the full realization of engineering estimates for these end uses. However, the realization rates for auxiliaries and air conditioning are significantly below 1.0, indicating only partial realization. It is unlikely that these results suggest any rebound effects for these latter two end uses. It is considerably more probable that they imply data errors or inappropriate assumptions underlying the engineering estimates developed in the course of the evaluation. Engineering estimates of cooling loads, for instance, seemed considerably higher than suggested by the variation in actual consumption data across weather conditions. Of course, it must be kept in mind that data problems can have an appreciable effect on the results. For instance, it is possible that, in spite of the efforts expended in this project, some billing data relating to cooling loads (which could be separately metered) could be missing. It is also possible that some of the ventilation loads in billing data were ‘‘assigned’’ to other end uses by the regression analysis. On the other hand, this result could stem from oversizing of equipment. Nonetheless, the general story told by Table 1 appears to be a consistent one: engineering estimates of energy savings are generally clearly reflected in actual differences in energy consumption. Realization rates associated with project engineering estimates are generally quite high, with an overall average for all end uses of .96.

Table 1. *Estimated Realization Rates*

End Use	Realization Rate on Energy Savings
Interior Lighting	1.12
Space Heating	1.27
Air Conditioning	.51
Auxiliaries	.68
Exterior Lighting	1.26
Refrigeration	1.15
All End Uses	.96

Table 2 presents the realized savings estimates developed through the application of the realization rate model. The model was used to generate realized savings estimates at the site level, and these estimates were weighted and aggregated across sites. As indicated in Table 2 and illustrated in Figure 2, there is a fairly close correspondence between the engineering estimates developed by AEC and the realized savings estimates. Indeed, the estimate of total realized savings amounts to 96% of the total AEC engineering estimate. The primary shortfall of realized savings is found in the cooling and auxiliaries end uses. The realization rate analysis yields what can be called *gross* realized savings, in the sense that the estimates do not take into account the possibility of free-ridership.

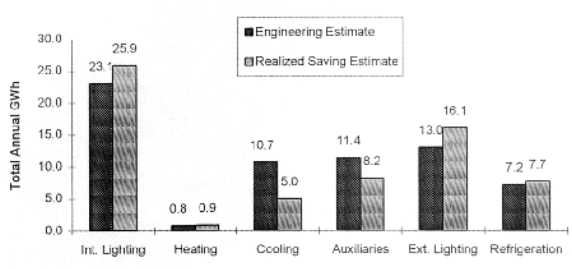
Estimation of Net-to-Gross Ratios

In the previous section, we discussed the estimation of gross realized savings associated with the Energy Smart Design Program. These gross savings relate to the DSM measures installed by participants in the program. Of course, some of these measures might have been adopted even in the absence of the program. This phenomenon is often termed the *free-rider effect*. In this section, we discuss the process of defining efficiency choices and estimating the net impacts of efficiency on these choices.

Table 2. Summary of Savings Estimates: All Utilities

End Use	Tracking System Estimates	Evaluation Engineering Estimates	Realized Savings
Int. Lighting		23,091,251	25,906,511
Ext. Lighting		12,950,071	8,351,856
Total Lighting		36,041,323	42,221,067
Envelope			0
HVAC		23,448,898	14,149,842
Space Heating		787,651	924,992
Cooling		11,000,498	5,044,786
Auxiliaries		11,660,749	8,180,062
Total Envel & HVAC		23,448,898	14,149,842
Water Heating		12,424	17,873
Refrigeration		7,163,297	7,744,710
Misc & Power		0	0
Total Other		7,175,721	7,762,583
ALL END USES	52,422,314	66,665,942	64,133,492

Figure 2. *Realized Savings by End Use (All Utilities)*



Defining and Estimating Overall End-Use Efficiency. Much of the literature in program evaluation concentrates on the effects of utility programs on the adoption of discrete DSM measures. This approach is sensible for the analysis of programs with purely prescriptive offerings, like high-efficiency air conditioning or compact fluorescent programs. However, new construction programs cover multiple end uses and a variety of DSM equipment and measures that impact each use. Satisfaction of code and (in many cases) adherence to program requirements may be accomplished on a performance, rather than a prescriptive, basis. A builder can adopt a wide variety of measures and qualify for participation. To provide a comprehensive assessment of program impacts on energy efficiency decisions, the analysis focused on several indicators of energy efficiency, rather than on the adoptions of discrete measures. Each efficiency index (EFF_{be}) is an estimate of proportional realized savings relative to the adjusted reference consumption for an end use e and building b:

$$EFF_{be} = \hat{\alpha}_e(X_b)\hat{\beta}_e[EEBASE_{be} - EEACTUAL_{be}]/[\hat{\alpha}_e(X_b)EEBASE_{be}]$$

The numerator of this index represents realized savings, while the denominator reflects adjusted reference consumption. Table 3 presents estimates of the efficiency index based on the engineering results and realization rates. Three comments are in order with respect to these results. First, efficiency levels may seem low relative to expectations, but this is at least partly due to the choice of reference values for the engineering analysis. Market standard practices were used as reference points for some building features, while code was used for others. In one case (outdoor lighting) efficiency indices are high. This is because the reference point for HID lighting was set at mercury vapor, which is probably less efficient than standard practice. Second, note that participants are defined with respect to any and all program activities. That is, a site receiving only an incentive for lighting is considered a participant for the purposes of other non-lighting efficiency comparisons as well. This is necessary to reflect the interrelated nature of DSM decisions affecting various end uses, a phenomenon that is critical when performance paths of compliance are available. Third, it should be recognized that efficiency indices were defined identically for participants and nonparticipants, with nonpar-

Table 3. Participant and Non-Participant Efficiencies

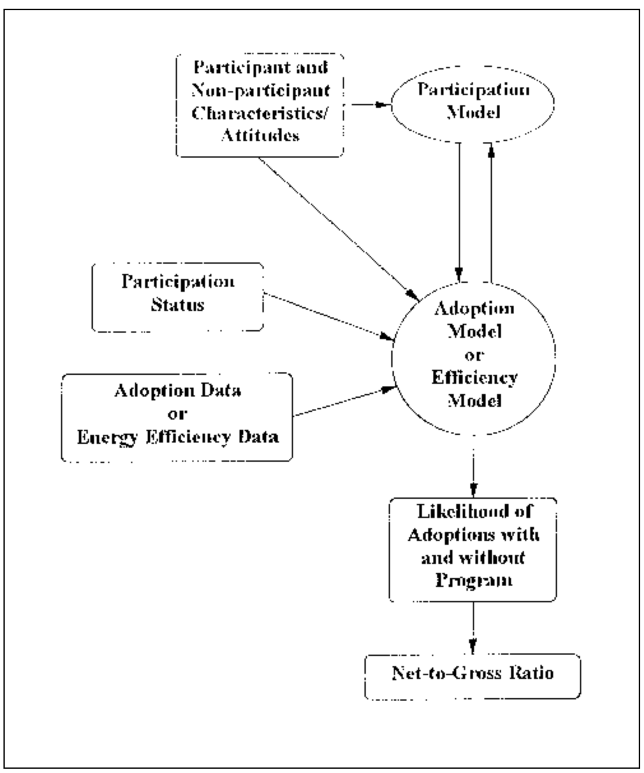
End Use	EFFICIENCY RATIOS: Participants	EFFICIENCY RATIOS: Non-Participants	Implied Net-to-Gross Ratios
Interior Lighting	.221	.149	.33
HVAC	.241	.134	.44
Exterior Lighting	.483	.432	.11
Refrigeration	.176	.209	−.16

ticipant indices derived from engineering analyses and estimated realization rates and adjustment coefficients.

Simple Comparisons of Participant and Nonparticipant Efficiency Indices. In general, the differences in efficiency between participants and nonparticipants is smaller than expected. If we were to use these simple comparisons to develop net-to-gross ratios (a common practice we do *not* recommend), they would yield the estimates in the last column of Table 3. These estimates may reflect strong free ridership; however, they may also indicate market transformation effects on non-participant behavior. To some extent, they may also reflect distortions caused by basic differences between participants and nonparticipants. Simple comparisons of efficiency levels do not control for these differences, and are thus highly suspect. The efficiency choice modeling approach described below mitigates these problems.

Modeling Efficiency Choices. The ultimate goal of the efficiency choice analysis is to estimate the *net* level of DSM adoptions actually attributable to the program (i.e., net of the free-rider effect). Conceptually, this entails comparing observed adoptions by program participants to the levels that would have occurred for these same participants without the program. Insofar as the latter levels of adoptions are not directly observable, they were estimated through the development of a statistical model of customer efficiency choice behavior. The model is illustrated in Figure 3.

Figure 3. *Efficiency Choice Model*



The model relates the adoption of energy efficiency to program participation, salient characteristics of the site and features of those making efficiency decisions. The model is estimated using information for a sample of participants and nonparticipants, and then used to simulate the net impact of the program, given the characteristics of participating sites. The ratio of net attributable adoptions to gross participant adoptions is called the *net-to-gross ratio*. The general algebraic form of the efficiency model used for the net-to-gross analysis is:

$$PARTt_b = f(EFF_{be}, INCENT_b, MARKET_b, SITE_b, DECISION_b)$$

$$EFF_{be} = g_e(PART_b, INCENT_b, MARKET_b, SITE_b, DECISION_b)$$

where

EFF_{be} is the efficiency level for end use e in building b,
$PART_b$ is a binary variable indicating participation in the New Construction Program,
$INCENT_b$ is a variable representing the incentive rate facing building b,
$MARKET_b$ is a set of market conditions facing building b,
$SITE_b$ is a set of site characteristics, and
$DECISION_b$ is a set of features relating to decision-making at the site.

The model used to analyze the impact of the Energy Smart Design Program was designed to include a reasonably large number of factors thought to affect efficiency decisions. These factors included the following:

- general building features like owner-occupancy, occupancy by private firms, chains and franchises, built-to-suit, building size, number of floors, window percentage, and a set of building category dummies;
- HVAC system features like the feasibility of variable air volume systems, the presence of electric space heat and air conditioning, the proportion of conditioned space with energy management systems, and the viability of heat pump applications;
- Lighting requirements (lumens per square foot) and the use of HID interior lighting;
- Weather conditions in the form of annual heating and cooling degree-days;
- Two indicators of the degree of code enforcement, one for lighting densities and one for U values, based on percentage of buildings satisfying codes relating to these attributes;

- program features, principally the incentive rate and an indicator of design-only program participation; and

- a set of binary variables representing the individual service areas.

One of the possible weaknesses of the Energy Smart Design efficiency model was that it contained no direct information on the attitudes of decision makers. It was decided early in the project that no decision-maker survey would be conducted. This decision was made partly on the basis of economics (like other evaluations, this project was faced with a limited budget). It was also based on practical considerations relating to the difficulties of collecting attitudinal data. Typically, new construction decisions involve decision teams. Members of the teams include both decision makers and decision influencers. Further, eventual building occupants rarely know about the decision-making process or criteria that were used in the design and construction phase. This makes it difficult to develop data relevant to modeling the net impact of programs on DSM measure adoptions.

Model Estimation. The participation equation and a set of efficiency equations can be estimated using data on efficiency choices, site features, decision-maker characteristics, a binary participation variable, and the factors affecting participation. Because of endogeneity of program participation and self-selection of the participants and nonparticipants, the simple empirical association of participation and adoptions will give a biased estimate of the effect of the former on the latter. The net effect of participating in the program is defined as the difference in the expected efficiency of participants who participated in the program and participants who would not have participated in the program had the program not existed.

To mitigate the presence of self-selection bias, three approaches can be used:

- First, a self-selection correction term (an inverse Mills Ratio) can be included in the efficiency equation. This term is a function of the predicted probability of participation, which is derived from the estimated reduced-form equation for the participation decision. This method is typically attributed to Heckman (1976).

- Second, the efficiency/participation model can be estimated using two-stage least squares, thus dealing with the simultaneous equation bias inherent in the application of ordinary least squares. This approach is often attributed to Hartman (1988). A similar approach involving nonlinear least squares with instruments is proposed by Train (1994).

- Third, the efficiency/participation model can be estimated simultaneously using Full Information Maximum Likelihood (FIML) estimation. This approach, developed by Wang (1994), is more efficient than the two-stage approach, but also mitigates simultaneous equation bias.

While the literature on self-selection has not yet yielded a clear consensus on the appropriate means of dealing with this problem in program evaluation, two impressions can be advanced. First, it is fairly clear that the Mills Ratio approach is often misused in the literature. Train (1994) argues, for instance, that the Mills Ratio best fits the case where the factors affecting the dependent variable of a regression (say, energy consumption) are affected by some of the same factors that affect a binary regressor (say, a binary participation variable). Train contends that the self-selection inherent in the evaluation of DSM programs occurs because the predisposition to adopt a measure affects the decision to adopt that measure, and that this is a different situation altogether. However, the Mills Ratio approach can still be justified for this study if we assume a certain structure of self-selection. Train argues that ‘‘the inverse Mills ratio is designed to handle a situation in which unobserved variables that affect the level of the dependent variable in a regression also affect a discrete choice (p 433).’’ If we interpret the dependent variable in the regression as the efficiency index, and if we characterize self-selection as phrased in this quotation, it would seem that the approach could be used. On the other hand, the Mills Ratio method has two additional practical problems. The first is that the self-selection correction term may be linearly correlated with the binary participation indicator ($PART_b$), and this can lead to the confounding of participation effects. The second problem is that its two-stage nature leads to inefficiency of parameter estimates, since efficiency and program participation are actually decided simultaneously.

Second, Train’s argument for the use of either the Hartman substitution approach or nonlinear least squares with instruments is compelling, given his perception of self-selection as a case of simultaneous determination of participation and adoption. However, neither two-stage method provides efficient and robust estimates. A poorly fitted probability of program participation may lead to an implausible estimate of program impact. The Wang FIML method, which entails the maximization of a log likelihood function for the two-equation simultaneous system, provides an efficient and consistent estimate of net program impacts.

Both the Heckman two-stage method and the FIML method were used in this project. The estimated participation and efficiency models are available on request from the authors. In what follows, we focus on key efficiency model results for interior lighting and HVAC.

Estimation Results. The estimated efficiency models offer some interesting findings with respect to the five programs being evaluated. While these findings may not transfer to other programs, they are nonetheless presented below as examples of the kinds of insights that can be developed with an efficiency modeling approach.

The estimated interior lighting efficiency model suggests several findings.

- Chains and franchises generally choose lower lighting efficiency levels than single-site establishments.

- The presence of interior HID lighting is associated with higher efficiency improvements (relative to the reference case).

- Large retail establishments tend to choose higher lighting efficiency levels.

- Both participant and non-participant lighting efficiencies tend to be lower in Idaho Power's service area, all other factors considered. This probably stems from the lack of new construction building standards in Idaho, as well as the lack of financial incentives for lighting efficiency.

- Design assistance has a significant impact on the choice of lighting efficiency, even in the absence of incentives. However, the influence of participation is significantly affected by the level of the incentive being offered under the program.

The estimated HVAC efficiency model also offers some interesting implications.

- Normal weather conditions (particularly heating degree days) affect the choice of HVAC efficiency.

- HVAC efficiency tends to be higher when variable air volume systems are feasible, presumably because these systems tend to be large and relatively efficient.

- The presence of an EMCS is typically associated with higher overall HVAC efficiency relative to the reference case.

- The viability of heat pumps has a significant effect on HVAC efficiency. Insofar as nonparticipants tend to have higher heat pump viability, the use of this term in the model controls for this phenomenon and increases the estimated impact of the program on HVAC efficiencies.

- HVAC efficiencies tend to be lower in the Idaho Power service area. Again, this is probably due to the lack of building codes in Idaho.

- The influence of program participation is significant in most versions of the HVAC efficiency model. However, the use of financial incentives does not appear to significantly increase this program impact. This latter result may stem from the relatively strong emphasis of the Idaho Power program on HVAC measures. While end-use estimates were not available from the IPC tracking system, the project estimates developed by AEC suggest a much higher HVAC savings share for IPC than for other utilities. In a sense, then, the regression is confounding apparent variations in program emphases with differences in incentive levels between IPC and the other utilities.

Computing Net-to-Gross Ratios

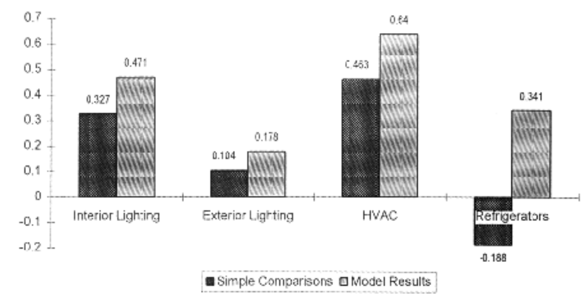
Once the efficiency models were estimated, they were used to assess the net impact of program participation on efficiency levels for each site. Based on these estimates, a set of net-to-gross ratios was computed for each service area. As shown in Table 4 and Figure 4, these estimates are considerably higher than the values obtained from the simple comparisons of participant and non-participant efficiency levels. This difference results from the model's capability to control for other factors affecting efficiency levels. Some of these factors proved to be important. For instance, participants tend to be considerably larger than nonparticipants and this makes installations of heat pumps less applicable, since larger buildings tend toward central systems with zone-level resistance heating.

The results suggest that overall net program savings for all utilities amount to over 25 annual GWh. This is 40% of the gross realized savings estimated in the course of the

Table 4. *Model-Based Net-to-Gross Ratios, All Utilities*

	NET-TO GROSS RATIOS		
End Use	Heckman	FIML	Average
Interior Lighting	.462	.480	.471
HVAC (comb)	.669	.610	.640
Exterior Lighting	.201	.155	.178
Refrigeration	.447	.235	.341

Figure 4. Comparison of Net-to-Gross Ratios (Model-Based vs. Simple)



evaluation using AEC engineering estimates and the results of the realization rate analysis. (This share is indicated by the unitalicized fraction at the bottom of the fifth column of Table 5.) However, net savings are almost 50% of the savings indicated by the tracking system estimates. (Net realized savings as a proportion of tracking system savings are indicated by the italicized fraction at the bottom of the fifth column.)

SUMMARY AND CONCLUSIONS

This paper describes the evaluation of a set of new construction programs operated by five Northwest utilities. The project was cofunded by the utilities and EPRI under a tailored collaboration. The project was intended to develop and apply a comprehensive and rigorous method of evaluating the gross and net impacts of these programs. The methodology developed for the evaluation consisted of four major steps:

- First, on-site data were collected for a sample of participants and nonparticipants.

- Second, engineering estimates of site usage were developed at the end-use level using DOE-2. Estimates were constructed under two scenarios: an as-built assumption and a reference scenario based on a mix of code and common practice. Differences in these as-built and reference estimates for participants can be interpreted as engineering estimates of gross savings.

- A realization rate model was developed to reconcile engineering estimates of savings with observed differ-

Table 5. *Summary of Net Savings Estimates, All Utilities*

End Use	Tracking System Estimates	Evaluation Engineering Estimates	Realized Savings Estimate	Net-to-Gross Ratio	Net Program Savings
Int. Lighting		23,091,251	25,906,511	0.471	12,201,967
Ext. Lighting		12,950,071	8,351,856	0.178	1,486,630
Total Lighting		36,041,323	42,221,067	0.324	13,688,597
HVAC		23,448,898	14,149,842	0.000	0
Space Heating		787,651	924,992	0.000	0
Cooling		11,000,498	5,044,786	0.000	0
Auxiliaries		11,660,749	8,180,062	0.000	0
HVAC & Envel		23,448,898	14,149,842	0.640	9,055,899
Water Heating		12,424	17,873	0.670	11,902
Refrigeration		7,163,297	7,744,710	0.341	2,640,946
Total Other		7,175,721	7,762,583	0.33	2,652,848
ALL END USES	52,835,826	66,665,942	64,133,492	0.396	25,397,344
				0.481	

ences in usage at sites with different levels of energy efficiency. The model yielded realization rates for each end use and site, and these rates were applied to the engineering estimates of savings to derive realized gross savings.

- Estimates of gross realized savings were then transformed into net savings through the application of a set of the net-to-gross ratios reflecting the portion of realized savings actually attributable to the program. The derivation of net-to-gross ratios was based on the analysis of efficiency indices encompassing the percentage impacts of all installed measures on the end-use consumption level. Several means of developing net-to-gross ratios using these efficiency indices were identified, including simple comparisons as well as three statistical approaches designed to mitigate self-selection bias. Two very important caveats should be offered with respect to these ratios. First, it should be understood that no matter how sophisticated the statistical analysis, the estimates may be biased downward by the presence of market transformation effects. What looks like free ridership may simply reflect the impacts of programs on non-participant behavior. To the extent that these effects were transmitted by *past* programs, they will not bias the estimates of free ridership, given that past programs are given and no longer under evaluation. However, if the program year under evaluation has influenced current non-participant behavior, the bias could be serious. Second, it should be noted that some of the current programs may help to transform the *future* market for efficiency. While this phenomenon will not bias estimates of current free ridership based on efficiency modeling, this transformation effect should be considered as a separate program benefit not covered by the methodology discussed in this paper.

The second project objective was to apply the methodology to the evaluation of the five programs under consideration. This application yielded the following general results:

- First, engineering estimates constructed during the course of the project differed fairly substantially from those developed by the participating utilities in the course of program implementation. In general, the evaluation estimates were larger than the utility estimates, partly because they covered all efficiency improvements, not just those claimed by the programs. However, other differences were also found, and seemed to be traceable to variations in assumptions and engineering algorithms.

- The realization rate analysis suggested that the engineering estimates of savings constructed during the evaluation were largely realized in the form of reductions in energy usage relative to the reference case. The overall realization rate for all end uses was 96%. However, there was some evidence that the engineering estimates of air conditioning and auxiliaries were overstated, based on the low realization rates on these end uses.

- The efficiency analysis yielded estimates of end-use net-to-gross ratios ranging from .18 for outdoor lighting to .64 for HVAC. The overall average net-to-gross ratio was estimated to be .39. In general, these estimates, which were based on the use of efficiency modeling techniques designed to mitigate self-selection bias, were higher than those derived from simple comparisons of participants and nonparticipant efficiency levels.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance of Thomas A. Mayer.

ENDNOTE

1. Note that other firms on the evaluation team included Xenergy, Architectural Energy Corporation, and Portland Energy Conservation, Inc.

REFERENCES

Atherton, T., C. Chen, R. Ridge, and K. Train. 1994. ''A Model of Appliance Efficiency Choice Using Stated and Revealed Preference Data.'' *In Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*, Volume 7. Washington D.C.: Amercian Council for an Energy-Efficient Economy.

Cardell, S. 1989. ''Extensions of Multinomial Logit: the Hedonic Demand Model, the Non-independent Logit Model, and the Ranked Logit Model.'' Ph.D. dissertation, Harvard University.

Greene, W. 1993. *Econometric Analysis*. New York, NY: Macmillan.

Griffin, J. 1993. ''Methodological Advances in Energy Modeling: 1970-90.'' *The Energy Journal* 14(1):111-124.

Hartman, R. 1988. *Review of Economics and Statistics, Volume 70.*

Hausman, A. 1986. ''Specification and Estimation of Simultaneous Equation Models,'' *Handbook of Econometrics*. Vol. III. New York: North-Holland.

Heckman, J. 1976. ''The Common Structure of Statistical Models of Truncation Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models,'' *Annals of Economic and Social Measurement*, 5/4.

Heckman, J. 1994. ''Evaluating the Impact of Training on the Earnings and Labor Force Status of Young Women: Better Data Help a Lot.'' Working Paper, University of Chicago.

Jacobs, P. and D. Roberts. 1995. ''Development of a Simplified Process to Estimate HVAC Energy Consumption for a Commercial New Construction Evaluation.'' *In Proceedings of Energy Program Evaluation: Uses, Methods, and Results,* 5C: 571-78. Chicago, IL: National Energy Program Evaluation Conference.

Maddala, S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge University Press.

Schiffman, D. 1994. ''A Monte Carlo Based Comparison of Techniques for Measuring the Energy Impacts of Demand-Side Management Programs.'' *In Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings,* Volume 7. Washington D.C.: Amercian Council for an Energy-Efficient Economy.

Train, K. 1986. *Qualitative Choice Analysis.* Cambridge, MA: MIT Press.

Train, K. 1988. ''Incentive for Energy Conservation in the Commercial and Industrial Sectors.'' *The Energy Journal* 9(3): 111-124.

Train, K. 1994. ''Estimation of Net Savings from *Energy* Conservation Programs.'' *Energy* vol. 19, No.

Wang, B. 1994. ''Maximum Likelihood Estimation with Sample Selection.'' Ph.D. Dissertation, Washington State University.