Reaching for High-Hanging Fruit: Cost-Effective Measurement of Demand Side Bids

Miriam L. Goldberg, XENERGY Inc.
Michael Stockard, TU Electric
Wm. Stevens Taber Jr., Princeton Development Corporation

Abstract

DSM bidding programs have historically attracted bids with high profit potential: those (1) serving large customers with homogenous usage, and (2) implementing short-payback measures (primarily lighting and drives). Other, less lucrative customer segments have been underserved by DSM bidding programs. Some utilities have attempted to reward "heroic" bids through qualitative scoring, but the viability of such bids has been constrained by prohibitively high measurement costs.

Frequently, these high measurement costs are driven by a requirement that a fixed precision criterion (such as 90/10 accuracy) be met for each technology at each treated facility. Applying the same precision standard at a more aggregated level can dramatically reduce the monitoring costs. However, defining such aggregate criteria in detail is not trivial.

The challenge is to develop measurement standards that (1) satisfy concerns of program sponsors and administrators regarding the validity of bidders' savings claims; and (2) keep total costs, including monitoring requirements, economically viable for bidders.

This paper describes the development of a mutually acceptable basis for defining precision requirements, and the practical application of statistical design principles to meet them. The authors examine sampling and measurement protocols that substantially reduce monitoring costs for bids with (1) a broad range of energy efficiency measures, (2) heterogeneous usage patterns, (3) high unit monitoring costs, and (4) low returns on investment.

The approaches are applicable to most demand side bids. Establishing cost-effective monitoring and verification standards will help performance contractors reach beyond the low-hanging fruit to provide more comprehensive energy efficiency packages. The value of these approaches will increase with the cost pressures of a deregulated, fully competitive electric industry.

Introduction

Energy efficiency in the electric industry first became an important resource during the 1980s, when the California Collaborative and similar initiatives fostered integrated resource planning and utility investments in efficiency. However, as electric utility restructuring has proceeded, energy efficiency has become less important as a resource acquisition activity and more difficult to justify politically. In many restructured markets, energy efficiency has been "left on its own" to compete directly against supply. In general, energy efficiency has not competed well under these circumstances for the following reasons:

1. Many of the historically important market barriers to investments in energy efficiency, which the utility resource acquisition programs were intended to overcome, remain in full force in a restructured industry. While past resource acquisition activities of utilities have reduced some of these barriers, some are less malleable and have not to date been reduced in a lasting manner.

2. Industry restructuring has brought the expectation of retail price volatility and lower prices, both of which make investments in energy efficiency, which tend to be fixed and long-term, riskier and less attractive.

One response to these changed circumstances, both by policy makers and practitioners of energy efficiency, has been an increased emphasis on performance contracting. Performance contracting offers several advantages, both as a market transformation tool and as a resource acquisition activity, which increase as the restructured market for electricity emerges and matures. These advantages include:

1. Performance contracting offers a natural opportunity to relieve the retail customer of the risks associated with volatile electricity prices and other restructured industry conditions.
2. If ratepayer-funded incentives for energy efficiency are phased out gradually as restructuring is implemented, incentive payments to performance contractors can be gradually reduced, weaning the performance contacting industry off of ratepayer subsidies as it becomes self-sufficient.
3. Performance contracting transfers the onus of overcoming market barriers from the retail customer, which generally faces a steep learning curve, to the experienced energy service professional, and in so doing, nurtures a permanent infrastructure of service providers who are adept at accomplishing projects. Esteves (1997) and Eto and Schlegel (1996) discuss this issue further.

Integral to the success of performance contracting is reliable, verifiable measurement and verification. Past attempts to implement the performance contracting business model have been frustrated, in part, by measurement protocols which relied to too great an extent on engineering judgments. This has led to customer dissatisfaction and disputes with service providers over the actual amount of performance. This problem has been partially overcome by the promulgation of the IPMVP (1997), which serves as a reliable and credible "meter" on saved energy.

However, sampling protocols that have typically been used in measuring savings have tended to be expensive. Protocols have typically required that stipulated precision levels be met for each stratum within each measure type within each site. Since the sample size per stratum is very insensitive to the population size, measurement costs in this approach are almost linearly related to the number of strata.

This problem of high measurement and verification cost was overcome in the early years of performance contracting by selecting primarily the most lucrative projects for implementation. For example, lighting retrofits in supermarkets can tolerate inefficient sampling procedures because:

1. lighting savings are relatively simple to monitor and measure,
2. the population of fixtures in a supermarket or in a group of supermarkets is relatively homogeneous, leading to inherently simple stratification, and
3. the very high return on investment, due to large numbers of similar fixtures and long operating hours, yields a large enough profit margin to absorb high measurement costs.

This phenomenon was noted by Goldman and Kito (1994). They found that out of 170 utility DSM contracts surveyed, 90 percent were for commercial/industrial customers. By end use, 50 percent of the 170 were for lighting only and the remaining "comprehensive" programs were themselves primarily (70 percent to 100 percent) lighting.

We believe that this phenomenon of overcoming high measurement costs by the "brute force" of high profits will be of necessity short-lived. The following circumstances will force the performance contracting industry to become more cost-efficient:

1. The "inventory" of energy efficiency investments will be depleted of "low-hanging fruit" forcing practitioners to seek opportunities in non-lighting end uses and in small, heterogeneous facilities.
2. The declining and increasingly volatile price of electricity will reduce or threaten returns on energy efficiency investments forcing practitioners to reduce risk by reducing costs.
3. The declining availability of ratepayer-funded incentives to energy efficiency investments, at least in some markets, will reduce the effective returns and force practitioners to produce projects more cost-efficiently.
4. The maturation of the energy efficiency industry will bring more companies into competition with each other putting downward pressure on prices of energy efficiency services.

The authors of this paper present a sampling protocol that reduces, by an order of magnitude over traditional techniques, the cost of measurement and verification in the small commercial sector, where cost control is particularly important. The protocol described below evolved from several previous protocols. A standard M&V protocol for the TU Electric DSM bid program was developed by Schiller Associates, based on the one they had previously developed for Pacific Gas and Electric Company's (PG&E) PowerSaving Partners (PSP) Program. The PSP M&V procedures were built on a variety of earlier efforts. Princeton Development Corporation (PDC) together with XENERGY Inc. and ADM had developed a modification to the PSP protocol, which was accepted by PG&E for PDC to use in that program. (A similar approach, developed by Schiller Associates and Andrew Goett, is now used as the standard sampling technique for other PG&E performance-based contracts.) PDC's modified PSP M&V protocol was further refined by the authors and Schiller Associates for use by PDC in the TU program.

In the TU Electric program, a customer service program is being implemented covering both lighting and HVAC upgrades and both retrofit and operating services. The range of services offered under the program makes the inherent cost of unoptimized sampling prohibitively expensive.

As discussed by Goldberg (1996), the specific sampling and precision requirements for M&V should be based on a clear understanding of the monitoring objectives and their relative importance. Common reasons for conducting monitoring and verification include:

1. to determine the value of savings, as a basis for determining payments to the performance contractor
2. as a quality assurance tool, to induce a higher level of performance
3. to provide information for future decisions.

Once these objectives and priorities are determined, established statistical principles of optimal design can provide substantial efficiencies in M&V requirements to meet them.

The development of the M&V plan for the TU Electric program described here followed these guidelines. Through a series of negotiations, a plan was developed that took into account the need for reliable information at varying levels and provided that information as cost-effectively as possible.

Particularly important in the development of this plan was the recognition that equal levels of precision for each treated end use does not necessarily represent the most effective use of M&V investments. In particular, the sampling plan developed provides a finer level of precision for lighting than for HVAC savings, while maintaining the target accuracy for the program as a whole. In optimizing the M&V resource allocation across the program components, the much higher cost for HVAC monitoring is taken into account as well as the relative magnitudes of the savings contributed by each component. As a result, the target accuracy is achieved for the program as a whole at a substantially reduced cost compared to that for traditional M&V sampling requirements.

Program Description

TU Electric's Pilot DSM Solicitation Project

TU Electric is the largest electric utility in the state of Texas and provides electric service to approximately 5.9 million people. TU Electric's service territory is 600 miles wide and 250 miles in breadth covering the middle third of the state of Texas.

In 1996, TU Electric began the implementation of eight DSM contracts that were selected through a pilot DSM solicitation conducted in 1994. The eight DSM contracts are ten years in length and target commercial/industrial customers for the installation of DSM measures. The estimated demand and energy impacts of these contracts are 76 MW and 400,000 MWh annually. Prior to beginning any installation of the DSM measures, each contractor was required to submit a Facility Monitoring and Verification (M&V) Plan. This Plan was to follow the *DSM Contract Procedures Manual* that was provided to all contractors by TU Electric in 1997. TU Electric and their consultant developed this Procedures Manual.

The PDC Program

PDC's program provides a full service outsourcing program for lighting and HVAC including efficiency upgrades and long-term programmed maintenance. This program is designed to help participating customers to:

- reduce their utility bills,
- reduce their O&M costs,
- improve the appearance, functionality, and comfort of their facilities, and
- concentrate their manpower and capital resources on their core business.

The program is potentially a preliminary model for a value-added "branded" electricity program for application in a deregulated marketplace.

Target market. The primary target customer segment is retail automobile dealerships. Auto dealerships represent a little understood but important part of the retail sector. In Texas, there are about 450 new car dealerships. These establishments represent only 1.6 percent of the number of retail establishments, but generate 18.7 percent of retail dollar sales. A typical new car dealer is estimated to have lighting loads of between 10 and 85 kW and total loads of 20 to 150 kW. Thus, a program targeted to this market could be significant in terms of regional energy efficiency.

Efficiency Measures Provided. The contractor is implementing the following energy efficiency measures at all customer facilities participating in the program:

1. A comprehensive lighting retrofit and efficiency upgrade
2. Programmed maintenance of the lighting system to ensure persistence of the savings
3. Programmed maintenance of the HVAC system to capture the full efficiency potential of the equipment installed

The programmed maintenance services are included in the package at no additional cost. However, the contractor does not intend to install HVAC equipment efficiency upgrades. Thus, the HVAC measures consist of the indirect savings from lighting efficiency improvements plus the savings from equipment maintenance.

Because the package of efficiency measures is uniform across the program, the approach to measurement and verification can be simplified resulting in cost savings for this component of the contractor's operations. The contractor's M&V strategy is described below. We first describe the negotiation between the contractor and the utility that led to the final agreement on the M&V plan. We then describe the lighting and HVAC sampling procedures in more detail.

Approach to Measurement & Verification

Guideline M&V Requirements

The guideline Measurement and Verification Plan developed for TU's bid program (TU Electric and Schiller Associates, 1997) required monitoring of a specific number of points for each usage area type. The number of points required depended on the variability within each group and on the number of points present. These requirements were designed to provide $\pm$10 percent precision at 90 percent confidence (90/10 precision) for each group at each facility. Contractors were asked to prepare specific Monitoring and Verification Plans for their programs consistent with this general protocol.

Proposed Alternate Approach

In preparing the M&V plan for the auto dealership program, PDC proposed a more cost-effective approach. The contractor argued for an approach that would provide 90/10 precision for the auto dealer program as a whole. This approach was advocated on the basis of a general approach to M&V for DSM programs.

In earlier years, extensive investments were made in highly accurate and expensive metering data for DSM monitoring and verification. Over time, practitioners have increasingly come to recognize the merit of investing in M&V based on the level of accuracy needed for decisionmaking and associated regulatory rulings. For general utility DSM programs, utility incentives and cost recovery are typically based on the performance of a program or set of programs as a whole; program planning decisions and approvals are based on the performance of a technology type across the entire group of participants. For a bidding program, the payment to the contractor is based on the performance of the contractor's projects as a group. In this context, it therefore makes sense to specify accuracy requirements for the bidder's activities as a whole.

The essence of the contractor's argument was two-fold:

1. The intent of the precision standards was to provide reliable savings estimates at the program level, hence the 90/10 standard should be applied at this level.
2. The sampling requirements for program components should be specified according to statistical principles of optimal design. That is, M&V resources should be allocated as efficiently as possible to achieve the program-level 90/10 target.

Utility Response to the Alternate Proposal

The utility was receptive to the proposal to target sampling precision at the program level rather than at the level of individual facilities. However, the utility requested the contractor to address some specific concerns related to the proposed sample design. The utility also had some concerns about the specific monitoring and analysis procedures proposed for HVAC units. The sampling issues the contractor was asked to address included the following:

1. A 90/10 precision standard to be met separately for lighting and indirect HVAC savings, the two end uses targeted by the program, not just for the combined savings.
2. Clear specification of how the sampling would occur so that sampling points would be spread across participating customers, not clustered in a few projects.
3. Inclusion of explicit oversight steps by TU Electric.

4. Testing of the accuracy of the savings estimates at the end of each year of monitoring to assure that the design is achieving the targeted precision.
5. Specification of the minimum number of points that would be monitored in the first year, even though the total number and size of participating customers was unknown in advance.
6. Justification of the assumed levels of variability for each sampling stratum.
7. Use wherever possible of the sample allocation formulas provided in TU Electric's M&V Plan rather than presenting alternate equivalent formulas requiring another stage of review by a qualified professional.

These concerns were all understandable. Most of them were readily addressed with clarifications and examples. However, the question of the level at which the 90/10 standard would be applied, whether for each end use separately as specified by the utility, or for the combined program as requested by the contractor, remained a sticking point.

The utility's concern about the reliability of the HVAC savings estimates was natural. The HVAC component of the program was much less standard, and its associated savings less predictable, than the corresponding lighting component. However, the HVAC component of the program was projected to account for only 20 percent of the program savings. The contractor showed that applying a 90/10 standard to each end use separately, rather than to the program as a whole, would increase the monitoring cost by a factor of four. While the precision for the HVAC and lighting components would each be increased by the extra monitoring, the overall program precision would improve only from 10 percent to 8.2 percent (at 90 percent confidence). In terms of the program as a whole, greater precision improvement could be had at lower increased monitoring cost.

Key Points of the Final Agreement

The final M&V plan agreed to by both parties balanced the utility's concerns for the reliability of the savings estimates against the contractor's need to control M&V costs if the program was to operate at all. A two-phase monitoring plan was established with specific provisions for modifications to the plan in the second and subsequent years. Key points in the final plan included the following:

- *For the first year, the contractor committed to monitoring a minimum of 347 lighting points and 32 HVAC units.* The number of lighting points was designed to provide 90/10 precision separately for each of six usage strata, based on preliminary assumptions of the variability (coefficient of variation or CV) within each stratum. The number of HVAC points was roughly double the number originally proposed by the contractor. The original number were designed to provide 90/10 precision for the program as a whole when combined with 90/10 precision for each lighting stratum separately. The 32 points were anticipated to provide approximately 31 percent precision at 90 percent confidence for the HVAC component alone.
- *In the second and subsequent years, the sample will be allocated across the lighting and HVAC strata in such a way as to give 90/10 precision for the program as a whole.* Assumed variability (CV) for the later-year sample design will be based on the measurements from the prior year(s). The allocation between HVAC and lighting points will take into account the much greater cost per point for HVAC monitoring using standard statistical sample formulas for optimal allocation with unit costs varying by stratum.
- *The lighting sample is stratified by six usage categories.* Within each of these categories, the sample is drawn at random across all points in all participating facilities.
- *The HVAC sample is stratified by two initial condition categories.*

- *In each year of monitoring, the sample size in each lighting stratum and in each HVAC stratum will be at least four.*
- *Systematic sampling methods, described below, will be used to distribute the points in each stratum over participating customers.*

These agreements honor the general principle that 90/10 precision for the program as a whole satisfies the basic objectives of the monitoring. At the same time, the first year samples are larger than is expected to be needed to satisfy an overall 90/10 precision. This higher sampling rate provides better estimates of the coefficients of variation that will be the basis for future sample size estimates as well as providing protection against overly optimistic assumptions for these CV's for the first year. In addition, the larger sample sizes provide a broader base of experience on which to judge the less well established HVAC measures. The minimum requirement of four observations per stratum for all years of sampling provides some protection against individual anomalous values as well as providing a reasonable basis for variance estimation.

The two-phase sampling approach provides a means to control the costs of monitoring. At the same time, for a given total monitoring cost, the approach will provide higher precision for the program-level savings of interest compared to an approach where precision targets are set in each functional use area. Moreover, the approach provides the flexibility of redefining the sample allocation over time without driving up the costs of monitoring inordinately.

The lighting and HVAC sampling procedures are described further below.

Lighting Sample

The lighting measures included in this program are

- Retrofits of existing fixtures, lamps and/or ballasts with more energy efficient fixtures, lamps and/or ballasts.
- Delamping with or without the use of reflectors.

In the case of both these lighting efficiency measures, the retrofit reduces demand, but the operating hours for fixtures are the same pre- and post-retrofit. Thus, the monitoring required for lighting savings is runtime hours. A sample of lighting points must be monitored in each year for ongoing program payments. However, it is not necessary to monitor the same points before and after installation of efficient equipment.

Stratification. No stratification was considered necessary at the facility or site level. The population of sites for this program will be relatively homogeneous. Automobile dealers tend to have fairly uniform operating practices because of protocols established by the manufacturers which they represent and because of competitive pressures from other dealerships. In addition, the manufacturers promote uniform operating practices through their trade groups, and most dealers are located on auto rows and in similar clusters; for both these reasons, dealers tend to imitate each others' business styles.

Stratification is needed, however, within facilities because of the wide variation in hours of operating space use type. The contractor conducted a substantial amount of research into this facility type in the course of other development programs serving automobile dealerships. On the basis of this research, the space within sites for this program was divided into the following six strata:

1. **Interior Sales And Display**, including merchandise display, presentation and closing areas, sales management, transaction areas, merchandise preparation areas, appraisal areas, sales support facilities, and related employee facilities.

2. **Parts and Inventory**, including parts and inventory warehousing, shipping and receiving, inventory control, inventory processing, support facilities, and related employee facilities.
3. **Administration,** including management, clerical functions, records, MIS, training, and related employee facilities.
4. **Service,** including repair and maintenance bays, body shops, alignment facilities, specialized service areas, tool storage areas, service support facilities, and related employee areas.
5. **Continuous,** including lighting operating continuously for purposes of security, safety, and display.
6. **Exterior,** including lighting operating during nighttime hours of purposes of security, safety, and display.

This stratification reflects the basic functions of a retail automobile dealership and has been used by the contractor in its other automobile dealership programs with good results.

First-Year Sample Size. In the first year of monitoring, the sample is designed to provide 90/10 precision within each of the six strata. The sample size n required to obtain 90/10 precision for any group is given by the standard formula for simple random sampling.

Real-Time Sampling Protocol

A practical difficulty with a simple random sampling approach is that the contractor naturally wants to install meters as sites enter the program. Thus, it is necessary to begin selecting points for the sample before the total population of points is identified.

One strategy would be to allocate sample points to the first few sites that join the program and stop when the target number has been reached. The disadvantage of this approach is that points in later sites have no possibility of being in the sample. Formally, then, the later sites are not included in the population represented by the sample. That is, the selected sample is technically a simple random sample from the points in the early group, not from the set of all points in all participating sites.

To deal with this difficulty, the contractor determined at the outset a minimum total number of points N_{minh} anticipated for the population in each stratum h by the end of the year. A systematic sampling rate k_h in each stratum was then calculated as

$$k_h = N_{minh} / n_h$$

where n_h is the sample size determined by the simple random sampling formula.

For example, if a minimum of N_{min1} = 2,000 points were anticipated for stratum 1, and the target sample size was = 68, the rule would select every 29th point for metering: 2,000/68 = 29.4.

With this strategy, the target sample size can be met with what is essentially a simple random sample from the first 2,000 points included in the program. If the anticipated minimum N_{minh} is not achieved by the end of the year, the contractor will select a supplemental random sample from the full set of participants to complete the target sample size.

This procedure allows points to be selected for monitoring as intake visits are completed, while ensuring that the sample is distributed across both early and late participants. Points are selected at each site after the intake visit and before the retrofit installation, so that the monitoring equipment can be installed at the same time as the retrofit equipment.

Sample Precision

Table 1 shows the expected precision that will be achieved for the first year lighting sample based on CV's observed in similar work. The first year sample is designed to provide 90/10 precision for each usage group separately. The resulting precision for lighting as a whole is 3.8 percent at 90 percent confidence.

Table 1. Expected First-Year Lighting Sample

Stratum h	Mean energy savings (kWh/yr) $\bar{y}_h$	Std Dev. energy savings (kWh/yr) s_h	Observed CV_h	Population Count N_h	Assumed CV_h	Sample Size n_h*	Precision achieved (90 percent confidence)
Administration	309	150	0.485	13,408	0.05	68	
Continuous	465	5	0.011	10,626	0.15	7	
Exterior	62	33	0.532	6,073	0.50	68	
Interior Sales & Display	323	156	0.483	11,583	0.50	68	
Parts and Inventory	443	235	0.530	7,225	0.50	68	
Service	127	62	0.488	5,213	0.50	68	
Total	315			54,128		347	3.8%

* First-year sample size based on assumed CV. Calculated precision at 90 percent based on observed CV.

HVAC M&V Plan

General Approach

The monitoring and verification of HVAC savings is designed to provide verified savings for the combined effects of the indirect savings from lighting retrofits and the HVAC equipment maintenance. Verified savings are determined by metering a sample of HVAC sites before and after the lighting retrofits and HVAC maintenance are implemented. From the metered sample, the savings are calculated on a per-ton basis. The unit savings are then applied to the total installed HVAC capacity in the participating sites to estimate total savings.

Stratification

In the second and later years of the program, the sample for the HVAC component of the program will be designed to provide sufficient accuracy for this component that the overall precision level for all measures combined is 90/10. The first-year sample is larger than what is expected to be needed to achieve this target. However, the first-year sample is stratified in the same way, and the same type of analysis is planned. In both years, a random sample is taken from the pool of all HVAC units in participant buildings.

The HVAC sample is not stratified by space usage type. Many HVAC units will serve multiple usage areas. Some usage areas such as service are likely to have no HVAC system serving them.

The sample is stratified by the pre-maintenance condition of the HVAC unit. This condition is determined as part of the intake audit. This stratification is used because the level of savings associated with the maintenance measures will vary greatly depending on the initial condition. Dividing the units according to the initial condition results in more homogeneity within each group and is expected to result in a more efficient sample design. The sampling plan assumes stratification into two initial conditions, (1) well maintained and (2) poorly maintained. Field procedures included an objective protocol for classifying units into these two strata.

Another factor that will affect the level of savings is the size of the unit. We expect the size of units to be in the range of 5 to 25 tons in most cases. Because we will be normalizing by size in the analysis, as discussed further below, and the anticipated range of size variation is modest, we do not anticipate a need to stratify by size.

First-Year HVAC Sample

In the first year, per the negotiated agreement, the HVAC sample will consist of a total of 32 monitored units, 16 poorly maintained and 16 well maintained. To estimate the precision of this HVAC sample, the following assumptions are made:

a. The total savings from the HVAC component is about 20 percent of the total program savings. This assumption is based on the program requirement that nonlighting measures account for a minimum of 20 percent of the total program savings combined with TU's request that the claimed HVAC savings from this program not exceed 20 percent.
b. Within each HVAC stratum (1 and 2), the CV of (normalized) HVAC energy savings is 1.0.
c. The HVAC savings per ton, for indirect lighting effects and maintenance combined, is about twice as great for units in poor initial condition (stratum 2) as for units in good initial condition (stratum 1).
d. An equal number of units in good and in poor condition are treated in the first year. This assumption is made for convenience as a basis for developing a preliminary estimate of the HVAC sample precision. The actual precision will depend on the mix of good and poor units as well as on the actual CV within each of these strata.

Under these assumptions, the first-year HVAC sample by itself is calculated to have a precision of 31 percent as noted above.

Computing Program-Level Savings

For each year of monitoring, HVAC savings will be calculated on a normalized basis using stratified ratio estimation (Cochran, 1977). The calculation is as follows.

First, the savings per ton in each stratum h is calculated from the metered sample as

$$\hat{R}_h = \bar{y}_h / \bar{x}_h = \left(\frac{1}{n_h}\sum_j y_j\right) / \left(\frac{1}{n_h}\sum_j y_j\right)$$

where the summations are over all units in the metering sample in stratum h, and y_j and x_j, respectively, are the savings determined from analysis of metering data for unit j, and the capacity in tons for the unit. The total savings Y_h in each stratum is then calculated as the product of the savings per ton R_h and the known total tons X_h of units in the stratum:

$$\hat{Y}_h = \hat{R}_h X_h .$$

Total HVAC savings is then the sum of the stratum total Y_h over all strata.

$$\hat{Y}_T = \sum_h \hat{Y}_h = \hat{Y}_{good} + \hat{Y}_{poor}.$$

With this estimation method, the deviation used to compute the CV in each stratum h is not the usual standard deviation of savings within the stratum. Instead, the deviation is the root-mean-square deviation of savings values y_j from the ratio line.

As an alternative to using ratio estimation, we could deal with the size variation by stratifying the sample on size. Ratio estimation is often used as an alternative to stratification. That is, with the normalization provided by the ratio method, size stratification is not necessary. Avoiding the need to stratify by size is helpful, because it is difficult to anticipate what the likely distribution of sizes might be. As a result, it would be difficult to allocate the monitoring sample to the size categories in advance, particularly in the first year. In addition, with a relatively small total sample size, we could not have very many size strata. We believe that we will obtain better overall precision through the use of the ratio estimator than we would by stratified sampling with a simple mean estimator.

M&V Cost Savings

Table 2 compares the sample sizes and rough cost estimates for the optimized sample sizes with those that would be required under the general program guidelines that required 90/10 precision in each usage group. The costs per point are based on initial projections of monitoring equipment costs, the frequency of equipment rotation, and site labor costs. The actual unit costs under the program have run somewhat higher than those shown in the table, but the relative cost comparison is still meaningful.

Table 2. M&V Sample Sizes and Costs for the Optimized Plan and the General Guidelines

						Optimized for Total Program			Guideline Approach: 90/10 Required by Usage Area		
Component	Total Energy Savings (kWh/yr)	Mean Energy Savings (kWh/yr)	Population Number of Points N	CV	Unit Monitoring Cost	Sample Size n	Total Cost	Precision at 90 % Confidence	Sample Size n	Total Cost	Precision at 90 % Confidence
Lighting											
Administration	4,143,072	309	13,408	0.49	$40	24	$960	0.16	65	$2,600	0.10
Continuous	4,941,090	465	10,626	0.01	$40	4	$160	0.01	1	$40	0.02
Exterior	376,526	62	6,073	0.53	$40	4	$160	0.44	76	$3,040	0.10
Interior Sales & Display	3,741,309	323	11,583	0.48	$40	22	$880	0.17	63	$2,520	0.10
Parts and Inventory	3,200,675	443	7,225	0.53	$40	20	$800	0.19	76	$3,040	0.10
Service	662,051	127	5,213	0.49	$40	4	$160	0.40	65	$2,600	0.10
TOTAL LIGHTING	17,064,723		54,128			78	$3,120		346	$13,840	
HVAC											
Initial Condition Good	1,322,500	9,121	145	0.85	$500	4	$2,000	0.69	84	$42,000	0.10
Initial Condition Poor	2,990,000	17,086	175	1.05	$500	11	$5,500	0.50	111	$55,500	0.10
TOTAL HVAC	4,312,500		320			15	7,500		195	$97,500	
TOTAL PROGRAM	21,377,223	393	54,448			93	$21,240	0.10	1082	$222,680	0.03

For the HVAC measures, the sample sizes for the guideline approach are understated, because they assume that the two sampling strata used in the optimized design would also be acceptable under the base approach. However, these strata are not usage groups, as required in the guideline method, but condition-based categories. If 90/10 precision were required for HVAC by the same usage categories as defined for lighting, much higher sample sizes would be required with much higher corresponding costs.

Even with this understatement of the base plan M&V costs, the costs are lower by more than a factor of 10 with the sample optimized for program-level savings. The program-level optimization plan does not provide 90/10 precision for each usage group, but does provide this precision level for the program as a whole. When program-level savings are of primary interest, this type of precision standard is appropriate, and the optimized design meets this standard at substantial cost savings. As a result, measures that might otherwise be unaffordable simply on the basis of M&V requirements may become a practical component of an energy-efficiency package.

Conclusions

Application of high precision requirements at the individual building level for programs designed to serve an entire service territory can result in prohibitively high monitoring and verification costs and discourage implementation of any but the most routine measures. This paper has described how the adoption of a program-level perspective can make forays into other kinds of offerings attractive to performance contracting vendors while still providing purchasers of performance contracting services the assurance of reliably estimated savings.

The sampling approach described here offered efficiencies in several ways.

1. A two-stage approach is used where the first-year samples are prescribed to assure good information for designing the second- and later-year samples.

2. With an emphasis on precision for the program as a whole, lower precision is accepted for program components that account for relatively small portions of the total savings.
3. The higher cost of sampling HVAC units compared with lighting is taken into account in the sample design allowing the best overall precision to be obtained for a given investment in M&V.
4. A real-time sampling protocol for lighting points allows monitoring devices to be installed and retrieved as each site is visited for pre-installation inspections and for efficiency measure installation. This approach avoids the need for re-visits and also allows monitoring to be distributed throughout the period of measure implementation.
5. For HVAC, ratio estimation is expected to provide better precision for a given sample size. This approach also avoids the need for stratification by size, which could be difficult to implement effectively in this context.

Developing an approach that was agreeable to both the utility and the contractor required attention by each party to the concerns of the other. The principles and procedures developed in this setting point the way to expanded opportunities for successful performance contracting in the varied forms such programs may take in the future. With M&V requirements that simultaneously satisfy program managers' needs for reliable savings estimate and contractors' needs for cost containment, performance contractors can reach beyond the low-hanging fruit to get higher value and higher savings.

References

Cochran, W. G. 1977. *Sampling Techniques.* New York, NY: John Wiley & Sons, Inc.

Esteves, Richard. 1997. *Market Barriers and the Standard Performance Contract (SPC),* September 16, available at www.cbee.org.

Eto, J., R. Prahl, and J. Schlegel. 1996. *A Scoping Study on Energy Efficiency Market Transformation by California Utility DSM Programs*, July, available at www.cbee.org.

Goldberg, Miriam L. 1996. "Reasonable Doubts: Monitoring and Verification for Performance Contracting." *Proceedings of the 1996 ACEEE Summer Study on Energy Efficiency in Buildings.*

Goldman, Charles A. and Kito, M. S. 1994. "Demand-Side Bidding; Six Years Later...," *Proceedings of the 1994 ACEEE Summer Study on Energy Efficiency in Buildings.*

International Performance Monitoring and Verification Protocol. 1997. *Measurement and Verification (M&V) Guidelines for Federal Energy Projects.* Developed in conjunction with the North American Energy Measurement and Verification Protocol. (available on-line at http://eande.lbl.gov/CBS/femp/MVdoc.html).

Schiller Associates. 1994. *PowerSaving Partners Measurement & Verification Procedures Manual.* Oakland, CA: Schiller Associates.

TU Electric and Schiller Associates. 1997. DSM Contract Project Procedures Manual. Dallas, TX: TU Electric.