# Double Ratio Analysis: A New Tool for Cost-Effective Monitoring

***Roger L. Wright, RLW Analytics***
***Marvin Herowitz and Ilene Obstfeld, XENERGY Inc.***
***Susan Butler, Pacific Gas and Electric Company***

Double ratio estimation is a statistical technique for combining (1) program tracking information in the population, (2) on-site visits and engineering modeling in a large sample and (3) short-term monitoring in a small sample. The approach is one way to leverage expensive short-term monitoring with supporting information. This paper describes the methodology and discusses its use for lighting retrofits in the impact evaluation of PG&E's CIA program. The paper also discusses the difficulties of collecting pre/post information on a rigorous sampling basis and recommends several alternative approaches.

## Introduction

Double ratio estimation is one way to leverage expensive short-term monitoring with supporting information. This paper describes the methodology and discusses its use for lighting retrofits in the impact evaluation of Pacific Gas and Electric's (PG&E) Commercial, Industrial and Agricultural (CIA) Energy Efficiency Rebate program for 1992 participants. In the CIA evaluation, double ratio estimation was used to combine (1) program tracking information in the population, (2) on-site visits and engineering modeling in a large sample and (3) short-term before/after monitoring in a small sample. The paper discusses both the foundations of the double ratio methodology and the difficulties of pre/post sampling.

Double ratio estimation is a blend of double sampling (sometimes called two-phase sampling) with ratio estimation. The statistical foundations are discussed in several references (e.g., Cochran 1977, Sarndal et al. 1992, and Appendix C of Pacific Gas and Electric 1993). The application of double ratio estimation to DSM evaluation was first discussed in Townsley and Wright 1990, A more in-depth discussion of PG&E's CIA application is in Pacific Gas and Electric 1993.

## The PG&E CIA Rebate Program

The CIA Rebate program, which began in 1990, provides cash incentives to commercial, industrial and agricultural customers who install a wide range of energy efficiency measures. The program covers end uses such as lighting, HVAC, agricultural measures, motors, refrigeration, and industrial processes. The CIA program is comprised of two subprograms. In the Customized program, the customer receives a rebate that is directly related to the calculated energy savings for the retrofit. The Express program is a direct-rebate program that offers the customer a set dollar amount for specified equipment.

### Motivation for Double Ratio Estimation

Double ratio estimation was used, together with a more conventional billing analysis, to estimate the gross energy savings for the CIA program. The analysis statistically combined project-specific estimates of participant energy savings from three sources: (1) the CIA program tracking system which provided information from each rebate application, (2) engineering modeling carried out in a relatively large sample, and (3) short-term end-use metering in a small sample.

The double-ratio approach reflects the following assumptions. First, although program tracking system estimates of participant savings are essential to operating and monitoring DSM programs, their estimates are not, in and of themselves, sufficiently reliable for broad acceptance by policy makers. Second, engineering analysis and models of energy savings can often provide improved engineering estimates of DSM measure savings and useful information on the underlying parameters and assumptions used to plan and evaluate DSM programs. Finally, end-use and appliance monitoring has the potential to provide the most accurate information on how DSM measures affect energy consumption.

The goal was to estimate the CIA program's total gross savings with adequate reliability. The difficulty with employing monitoring without supporting information is that large samples are generally needed. Monitoring must be timed to measure all the major determinants of DSM savings in both the pre- and post installation periods. Monitoring must be maintained for a period of time representative of the conditions under which the DSM measures will operate over their lifetimes. Moreover, the accuracy of the measured results can be jeopardized by a host of other factors including calibration and accuracy problems, sampling problems among buildings and within buildings, interaction between the treated end use and other end uses, and the wide diversity of building types and measures in the target population of participants. In short, it is costly and time-consuming to do a good job of monitoring in a large sample.

For these reasons, monitoring studies traditionally have been limited to unique situations where it is cost-effective to collect high resolution data on a small number of customers and a narrow range of technologies. For large scale DSM programs such as CIA, with thousands of participants and many qualifying technologies, monitoring is often thought to be too limited and impractical to be used to provide reliable evidence of program impacts.

Engineering models come in many forms. Their common feature is that they combine participant-specific building equipment and end-use information with engineering algorithms to estimate energy savings from DSM measures. Engineering-model estimates of energy savings may be more accurate than program tracking system estimates. This is because simulations calculate end-use consumption in a more complex and realistic manner than the streamlined, non-interactive formulas that are often used on program applications, especially for programs such as the Express program. Also, they often use better site-specific information, typically requiring one or more surveys or inspections of each participant site that is modeled. However, the costs of engineering modeling are generally significantly lower than monitoring costs, and thus engineering sample sizes can be substantially larger than monitoring samples.

The double ratio estimation methodology attempts to exploit the strengths and avoid the weaknesses of each of these measurement techniques. It was designed to make efficient use of costly, high resolution monitoring data in a small sample by using these data to calibrate less costly, less precise engineering-model estimates of savings for a much larger sample of participants. The engineering-model estimates were used, in turn, to adjust the low cost, least accurate tracking system estimates of savings available for the entire population of 1992 program participants.

## The Sample Data

This paper uses the CIA study to illustrate the double ratio methodology. The samples consisted of 61 program participant sites that were monitored during their pre- and post installation periods and 282 program participant sites that received on-site inspection and had their energy use analyzed using an engineering model. The 61 monitored sites were a subsample of the 282 engineering-model sites so we could compare the modeled results to the monitored results for each of these sites. Table 1 shows the breakdown of sites by end use and program. For these sample sites as well as for the entire 1992 CIA participant population of 7,990 sites, tracking system estimates of energy savings were also available.

**Table 1.** Sample Sizes

| End Use | Program | Modeled Sites | Monitored Sites |
|---|---|---|---|
| Lighting | Expr | 96 | 16 |
| | Cust | 133 | 36 |
| HVAC | Expr | 25 | 0 |
| | Cust | 18 | 9 |
| Refrig | Expr | 4 | 0 |
| | Cust | 6 | 0 |
| Total | | 282 | 61 |

A two-phase sampling plan was developed for selecting the monitored and modeled sites from among the population of program participants. The sampling plan was stratified by program type, measure category and the tracking estimate of savings. Model-based statistical sampling techniques were used to construct the strata and allocate the sample optimally given the information that was available at the time the sampling plan was designed. The goal was to minimize the overall sampling error of the *combined* field monitoring and engineering model estimates of program energy savings for the available resources. The methods are described in Sarndal et al. 1992, Wright 1989, Appendix C of Pacific Gas and Electric 1993, and the next major section of this paper.

In practice, recruiting the samples for pre/post measurement was difficult and expensive. In the Express program it was nearly impossible to cost-effectively draw a sample for pre-retrofit measurement because the program provided no prenotification of intent to participate prior to the retrofit. Surprisingly, the recruiting problems were almost as difficult in the Customized program due to short lead times and inability to predict when projects would actually be implemented. The recruiting problems were especially serious for the pre-retrofit monitoring. To emphasize, the problem was not lack of customer cooperation but lack of information about upcoming projects.

Results from this study were reported in terms of realization rates. A realization rate is the ratio of the measured savings to the assumed savings in the tracking system. The double ratio analysis consisted of calculating two ratios, (1) the ratio between the average modeled savings and average tracking savings observed in the modeled sample and (2) the ratio between the average monitored savings and average modeled savings observed in the monitored sample. The two ratios were then multiplied to determine the overall realization rate.

To partially compensate for the recruiting problems, the final samples were post stratified and reweighted to the target populations. These weights were used to calculate the averages used in each of the ratios. The strength of association between the estimates of savings was calculated across the sites within each sample and used to evaluate the statistical precision of the double-ratio estimates.

## Lighting Results

As Table 1 shows, most of the sample sites were lighting retrofits so these will be used to illustrate the double ratio methodology. The period of monitoring varied from one lighting site to another, but on average, these sites were monitored for about 5 weeks before the retrofit and about 3 weeks afterwards. Table 2 contains the mean estimates of program energy and demand savings for lighting measures that were installed under the Express and Customized programs. The relative precision of the ratios of the sample estimates are provided within the table.

**Table 2.** Intermediate Results for Lighting Measures

| | 1st Phase Modeled to Tracking | | 2nd Phase Monitored to Modeled | |
|---|---|---|---|---|
| | **Expr** | **Cust** | **Expr** | **Cust** |
| **Sample Size** | 96 | 133 | 16 | 36 |
| **Short-term Monitoring** | | | | |
| Mean MWh Savings | -- | -- | 53.3 | 32.2 |
| Mean kW Savings | -- | -- | 10.5 | 5.2 |
| **Engineering Modeling** | | | | |
| Mean MWh Savings | 29.1 | 31.2 | 42.2 | 26.2 |
| Mean kW Savings | 5.3 | 5.7 | 8.1 | 4.8 |
| **Tracking System** | | | | |
| Mean MWh Savings | 34.4 | 51.2 | -- | -- |
| Mean kW Savings | 8.2 | 8.7 | -- | -- |
| **Results** | | | | |
| Ratio (MWh) Savings | 0.85 | 0.61 | 1.26 | 1.23 |
| Relative Precision | ±18% | ±8% | ±22% | ±21% |
| Ratio (MWh) Savings | 0.64 | 0.66 | 1.29 | 1.08 |
| Relative Precision | ±20% | ±10% | ±18% | ±17% |

The first-phase results indicate that the engineering model estimates in the Express program were 15% and 36% smaller than the corresponding program tracking estimates of the energy and demand savings. The relative precision of these estimates were ±18% and ±20%, respectively. This means, for example, that the 90% confidence interval for the true ratio between the engineering model estimates and tracking estimates is 0.85 x (1 ± 0.18 ) for energy savings. The second-phase results indicate that the short-term monitored estimates were 26% and 29% larger than the engineering model estimates of the energy and demand savings. The relative precision of these ratios were similar to those of the ratios of engineering modeled to program tracking savings. Analogous results were developed for the Customized program.

The three estimates of program-related savings were combined by multiplying together the first and second phase ratios. The results of these calculations are the realization rates displayed in Table 3. The error bounds were calculated at the 90% level of confidence using a procedure discussed in Appendix C of Pacific Gas and Electric 1993.

**Table 3.** Final Realization Rates for Lighting Measures

| | Expr | Cust | Total |
|---|---|---|---|
| **Energy (MWh)** | | | |
| Realization Rate | 1.07 | 0.75 | 0.93 |
| Error Bound | ±.42 | ±.09 | ±.26 |
| Relative Precision | ±39% | ±12% | ±28% |
| **Demand (kW)** | | | |
| Realization Rate | 0.83 | 0.71 | 0.8 |
| Error Bound | ±.26 | ±.11 | ±.23 |
| Relative Precision | ±31% | ±15% | ±29% |

## Caveats

The error bounds reported in Table 3 may underestimate the true uncertainty in the double-ratio results. Potentially the most serious problem was the recruiting difficulties in obtaining the before/after measurements. The double-ratio results rest strongly on valid sampling in both phases. If the sampling breaks down, it is difficult to quantify the extent of any resulting bias or to make an objective allowance for the additional uncertainty. The final section of this paper offers several suggestions for reducing sampling problems in future studies.

Other limitations should also be considered in interpreting these results. The variability from sampling the metered lighting circuits within the sites was not explicitly taken into account. However, this source of variation may be reflected in the total variation and therefore in the final precision. The process for determining the savings in the monitored sites may also not be as accurate as desirable. In particular, it may be appropriate to give additional attention to interaction effects between lighting and HVAC.

## Planning an Evaluation with Double-Ratio Estimation

This section will summarize various elements of the double-ratio methodology in greater detail. The emphasis is on tools for planning a new study similar to the lighting component of the CIA study.

### The Error Ratio

The *error ratio* is the key parameter in planning a project involving ratio estimation with optimal stratification. First consider the case of a one-phase sample design. The realization rate is to be estimated by calculating a single ratio relating the average measured savings to the average value of the tracking-estimates of savings for the sample.

The central issue is to choose the sample size needed for the required relative precision with ratio estimation. The expected relative precision and error bound were derived following the principles of statistical sampling adopted to ratio estimation with optimal model-based stratified sampling. In the one-phase case, the expected relative precision is primarily determined by the size of the sample and the variability in the population. For ratio estimation with a single sample of size $n$ the equation is:

$$rp = 1.645 \frac{er}{\sqrt{n}} \qquad (1)$$

Here *er* denotes a measure of population variability called the error ratio. For example, consider monitoring for lighting measures in the Customized Program. In this case, the error ratio was found to be about 37%. Wright 1992 defines the error ratio and describes several estimation techniques. We use the MBSS™ software for these calculations.

Given an error ratio of 37% and a sample size of 36, the expected relative precision would be approximately ±10%:

$$rp = 1.645 \frac{0.37}{\sqrt{36}} = \pm 0.10 \qquad (2)$$

In planning a project, the preceding equation can be easily solved for the required sample size, giving the equation:

$$n = 1.645 \left(\frac{er}{rp}\right)^2 \qquad (3)$$

The error ratio reflects the strength of the association between the two measures of savings that make up the realization rate. If the association is strong, the error ratio will be small and the relative precision will be correspondingly good. The association is strong if, after adjusting for the estimated realization rate, the tracking estimates generally give accurate project-by-project estimates of the measured savings observed in the sample.

Figure 1 illustrates an error ratio of 37% similar to that found in the Customized Lighting monitoring sample. The x variable is the engineering estimate of savings from the tracking system, The y variable is the corresponding savings measured from monitoring. Each point represents a particular project in the monitoring sample.

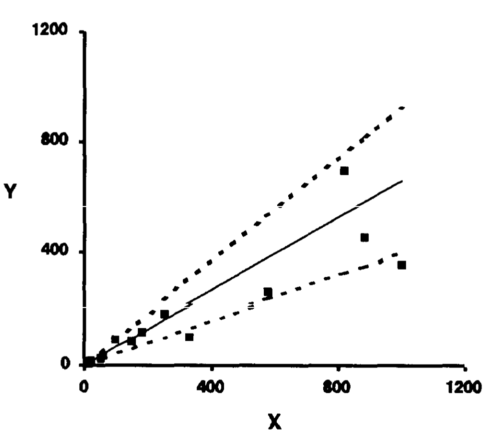


**Figure 1.** Example of an Error Ratio of 37%

The solid line is drawn from the origin through the point that represents the average value of the tracking estimates and the average value of the measured savings. The slope of this line is equal to the realization rate. This line represents the expected value of the measured savings for each project.

The measured savings of each project also has a standard deviation which represents its variability around the expected value, represented by the dashed lines in Figure 1. In most cases, the standard deviation increases with the expected value of the measured savings. In the simplest case, the standard deviation is directly proportional to the expected value. In this case, the error ratio is the standard deviation divided by the expected value. For example, if the standard deviation is equal to 37% of the expected value, then the error ratio is equal to 37%.

The error ratio, therefore, indicates the strength of support for the monitoring provided by the engineering estimates. If the error ratio is small, then the measured points will generally fall very close to the solid line. In this case, a small sample will provide a good estimate of the realization rate. In the extreme case that the error ratio is zero, even one measured point would be adequate.

Conversely, if the error ratio is large, then a large sample would be needed. For example, in the Lighting Express program, the error ratio was 110 %. In this case, a monitoring sample of about 200 projects would be needed for ±10% relative precision. In this situation, the monitoring sample might be reduced by developing improved engineering estimates using modeling in a larger supporting sample.

## Error Ratios in Double Sampling

In the double-sampling case, the expected relative precision is related to two pairs of parameters: (1) the error ratio and sample size in the larger first-phase modeled sample, denoted $er_1$ and $n_1$, and (2) the error ratio and sample size in the smaller second-phase monitored sample, denoted $er_2$ and $n_2$.

Here $er_1$ reflects the association between the tracking estimates of savings and the engineering model estimates of savings, while $er_2$ reflects the association between the engineering model estimates of savings and the monitoring estimates of savings. Then the expected relative precision is:

$$rp = 1.645 \sqrt{\frac{er_1^2}{n_1} + \frac{er_2^2}{n_2}} \qquad (4)$$

As an example, consider the Lighting Express case. In this example, the first-phase error ratio is 84% and the sample size is 96, while the second-phase error ratio is 46% and the sample size is 16. Using an optimally stratified sample design in each phase, the expected relative precision would be:

$$rp = 1.645 \sqrt{\frac{(84\%)^2}{96} + \frac{(46\%)^2}{16}} = \pm 23\% \qquad (5)$$

## Optimal Design in Double Sampling

A statistical study is said to be optimally designed if it is expected to provide a specified level of precision at the least cost. In the context of the CIA study, an optimal design provides the best allocation of resources within each phase and also between the first-phase modeling and the second-phase monitoring.

The preceding section described an equation for estimating the expected relative precision from estimates of the first- and second-phase error ratios, together with the sample sizes selected for the two phases. With some algebraic manipulation, together with suitable assumptions about the cost of the modeled and monitored samples, this equation can be used to design the optimal double-sampling study. The cost of the best double-sampling study can be compared to the cost of a one-phase approach that achieves the same expected precision using monitoring alone.

The results of this analysis will depend on the relative cost per unit for the modeled and monitored samples. The analysis presented in this section assumes that the engineering model work can be carried out for $1,000 per modeled project, while the field monitoring costs $10,000 per monitored project. Of course the actual costs will depend on the characteristics of each particular program as well as the approach to the modeling and monitoring data collection and analysis.

Table 4 is based on the error ratios for the energy savings (MWh) of lighting measures in the Express Program. Column four reports the sample sizes for the modeled and monitored samples under the optimal double-sampling approach. Column four also shows the sample size required to provide the same relative precision using a one-phase approach using monitoring alone. These results show that under the stated assumptions the best double-sampling experimental design requires 84 modeled and 15 monitored projects, while a one-phase approach giving the same ±25% expected precision requires 53 monitored projects.

Column five of Table 4 shows the costs. The effectiveness of the double-sampling strategy can be seen by comparing the total cost of the modeled and monitored samples with the total cost of the equivalent one-phase approach, in this example $234,000 versus $530,000.

The results shown in Table 5 are quite different. This analysis is based on the error ratios estimated for MWh savings in the Customized lighting program. In this situation, the best double-sampling approach is much more

**Table 4.** Optimal Sample Allocation for the Express Program

| | Error Ratio | Unit Cost | Sample Size | Total Cost |
|---|---|---|---|---|
| Modeled | 84% | $1,000 | 84 | $84,000 |
| Monitored | 46% | $10,000 | 15 | $150,000 |
| Total | | | | $234,000 |
| One-phase | 110% | $10,000 | 53 | $530,000 |

**Table 5.** Optimal Sample Allocation for the Customized Program

| | Error Ratio | Unit Cost | Sample Size | Total Cost |
|---|---|---|---|---|
| Modeled | 52% | $1,000 | 68 | $68,000 |
| Monitored | 78% | $10,000 | 32 | $320,000 |
| Total | | | | $388,000 |
| One-phase | 37% | $10,000 | 6 | $60,000 |

expensive than the one-phase approach using monitoring alone. This is because the error ratio in the monitored sample is actually larger than in the one-phase approach. In other words, the monitored savings were more weakly associated with the modeled estimates of savings than with the tracking estimates developed in the program itself.

A comparison of these two examples suggest that double-ratio estimation may be especially useful when the tracking estimates are weak as in the Express program. In the Customized program, on the other hand, the tracking estimates are strong and a one-phase approach to monitoring seems to be practical,

## Recommendations for Future Studies

A number of observations and recommendations can be made based on this experience.

The difficulty of collecting pre/post monitored data should not be underestimated. It is often hard to predict when a final decision to proceed with a project will be made. In many cases the project is implemented very quickly once the go-ahead is given. In other cases, the project is delayed beyond the period of the monitoring study. Unless information about most upcoming projects is available with adequate lead time and assurance that the projects will actually be implemented, it will be difficult to follow a pre/post sample design rigorously. Without valid sampling, the findings will be vulnerable to bias.

When valid pre/post sampling is impractical, alternative approaches should be considered. One possibility is post-retrofit monitoring. If post-retrofit information is adequate, most of the sampling problems can usually be avoided.

Post-retrofit monitoring should be considered when (1) the largest source of uncertainty concerns the hours of use of the new equipment, or (2) there is reliable information about pre-retrofit conditions, or (3) a standard retrofit is used as a baseline, i.e., the savings are calculated as the difference between the high-efficiency equipment installed under the program and standard equipment that would have been installed in the absence of the program. In commercial lighting retrofit programs, in particular, there is strong evidence that the hours of use are generally the same before and after the installation of more efficient lighting.

In some cases, it may be necessary to collect pre-retrofit information for a sample of upcoming projects. In this case it will generally be necessary to strike a balance between (a) insisting on longer lead times with greater uncertainty about implementation, and (b) developing a method to collect data quickly within shorter lead times. In most cases, it will be necessary to minimize the cost and duration of the pre-retrofit data-collection that is planned. For example, in a commercial lighting retrofit program, pre/post site inspections and spot wattage measurements might be combined with post-retrofit monitoring of hours of use. Alternatively, pre-retrofit conditions might be measured in a separate sample of upcoming projects.

Despite these problems, site-specific information about measure impacts can provide very valuable information, especially when valid sampling is used. The focus on realization rates and error ratios can provide more cost-effective sample designs that reflect the characteristics of each program. This approach can also strengthen the feedback loop between (1) evaluation and (2) program design and implementation. The double-sampling approach can add a second feedback loop between monitoring and modeling. Gradual improvement to the tracking estimates and models can be expected to yield smaller error ratios, improved precision, more effective programs, and higher customer satisfaction.

## References

Cochran, W. G., *Sampling Techniques,* 3rd Edition, Chapter 12, John Wiley, 1977.

Pacific Gas and Electric. 1993. Double Ratio Analysis Final Report, Xenergy Inc. and RLW Analytics, Inc.

Sarndal, C. E., B. Swensson, and J. Wretman, *Model Assisted Survey Sampling,* Springer-Verlag, 1992.

Townsley M., and R. L. Wright, “Measuring DSM Impacts: End-Use Metering and the Engineering Calibration Approach,” *End- Use Information and its Role in DSM*, The Fleming Group, 1990.

Wright, Roger L., Methods and Tools of Load Research, RLW Analytics, Inc. 1989.

Wright, Roger L., “Choosing sample sizes for sample designs with auxiliary information,” Workshop on Uses of Auxiliary Information in Surveys, Statistics Sweden, Orebro, 1992.