A COMBINED ENGINEERING AND DECISION-ANALYSIS METHODOLOGY FOR EVALUATING SPILLOVER AND FREE-RIDERSHIP IN PG&E'S 1995 INDUSTRIAL ENERGY EFFICIENCY PROGRAM

Lisa Reid, Pacific Gas and Electric Company, San Francisco, CA
Michael Baker, SBW Consulting, Inc., Seattle, WA
Richard Ridge, Ridge and Associates, Alameda, CA
Bing Tso, SBW Consulting, Inc., Seattle, WA

Introduction

An evaluation of PG&E's 1995 Industrial Energy Efficiency Program has been conducted to determine first-year gross and net kWh, kW and Therm impacts. The evaluation of net impact has been carried out using a combined engineering and decision-analysis methodology, which addresses the effects of both spillover and free-ridership, applied to more than 140 participant sites. This paper describes both the methodology and results of this net impact analysis.

Evaluations of net impact for industrial programs often rely on a participant self-report methodology, as it is difficult to implement statistical methods, such as discrete choice models or billing regression analysis of participant and control groups. Statistical methodologies are difficult to implement because of the great variability in industrial facility characteristics, with each facility representing nearly unique circumstances. Unfortunately, the only alternative to these statistically methods, participant self-report, is generally considered to be a biased technique. In recent years, a variety of internal consistency measurements have been added to the self-report methodology in an attempt to control measurement error and bias. This study adds some additional refinements which further reduce error due to omitted measurements and integrate an engineering analysis of customer-baseline savings (savings which would have occurred in the absence of the program incentive) to account for partial free-ridership.

One of the important measurements generally omitted from the self-report methodology is the influence of the utility's programs on trade allies who are the vendors of equipment and professional services needed to implement efficiency measures. Free-ridership can occur even if the participant had no prior intention of selecting efficient equipment. This can occur if the vendor has, independent of the utility's program, decided to recommend, promote or stock only equipment which qualifies for the program's incentive. In addition, participant spillover can occur by a similar mechanism. Vendors are influenced by the utility's program to endorse equipment which meets the programs qualifications. The customer may then install efficient equipment, without receiving a rebate or knowing of the program's influence on their selection. In this study, both of these omitted measurements have been addressed by introducing information from a survey of the vendors involved in the customer's efficiency choices.

Another weakness in previous self-report analyses has been the imposition of a binary choice assumption. It has been assumed that the customer chooses between the efficiency measure offered by the program and the no-action option. In this study an additional customer option is allowed through the incorporation of an engineering analysis of the customer-baseline alternative. In some cases the customer, in the absence of the program, would have installed equipment which was somewhat more efficient, but not as efficient as the equipment implemented under the program. By analyzing the savings associated with this "customer-baseline" it has been possible to quantify the effect of partial free-ridership. The engineering analysis of customer-baseline savings relies on information gathered through telephone and in-person interviews, on-site inspection of affected systems and the analysis of customer data which describes the operation and performance of the affected systems before and after the installation of efficiency measures.

Methodology

Sample Selection

The sample frame for this study was developed from PG&E's program database entries for the 1645 items associated with the 862 applications paid in 1995. Item-level entries described specific pieces of equipment for which a rebate was paid, e.g., a specific type of lighting fixture. These items were grouped into four domains: (1) Lighting (interior only), (2) Process, (3) HVAC, and (4) Miscellaneous (which includes refrigeration, motors, and exterior lighting). Item-level entries were aggregated to form a project-level sample frame. Projects were defined as the set of items listed on a single paid application that were assigned to the same end-use and control number (unique location identifier). The final sample frame contained 922 projects, associated with 703 control numbers. 167 of these projects were in the HVAC domain, 479 were in lighting, 182 were in miscellaneous, and 94 were in process.

Evaluations were completed for two samples. The first sample (the project-specific sample) provided a census of 70% of the program database kW, kWh, and therm savings for the light, process, and HVAC domains. The second sample (the verification sample) completed a census for the HVAC and process domains, since these domains contained fewer than 150 projects. For the lighting domain, the second sample contained enough projects to bring the total sample count for the domain up to 150.

Overview of Site Evaluation Process

Figure 1 provides an overview of the evaluation process for a sampled site. The first step in the process was to assign the site to a lead engineer. The lead engineer recruited the site by obtaining permission for the data collection activities necessary to evaluate all of the projects and spillover measures located at the site. Once a site was recruited the lead engineer was responsible for conducting a detailed review of the relevant application files and thoroughly understanding the projects present at the site.

Following recruitment and the detailed file review, the steps required to complete the evaluation work depended on the type of projects that were installed under the program. Three types were defined: (1) Customized Project-Specific (CPS) requiring project-specific estimates of gross savings and a customized evaluation of free-ridership; (2) Standard Project-Specific (SPS), requiring project-specific estimates of gross savings and a standardized evaluation of free-ridership (including an adjustment for partial free-ridership); (3) Verify-only (VO), requiring only a simply inspection of measures and the standardized evaluation of free-ridership (without an adjustment for partial free-ridership).

For both CPS and SPS projects, a site-specific evaluation plan was created which established the data collection and analysis methods to be used to determine gross and net savings. The plan also specified how savings were to be estimated for any spillover measures identified through a spillover survey. Data collection was conducted as specified in the plan, including an on-site survey and interviews with one or more members of the customer staff. For CPS projects, this included the administration of customized free-ridership related questions. In accordance with the site-specific plan, engineering and decision-analyses were completed to estimate gross and net savings for each project.

Definition of Full and Partial Free-ridership

Free-ridership occurs when customers receive rebates even though they would have implemented an efficiency measure without the rebate; hence, they are getting a "free ride" on the incentive program. In some cases, PG&E's programs motivate customers to replace equipment prior to the end of its useful life, an "early replacement" action. In other cases, the program motivates the customer to select more efficient equipment when replacing equipment that has reached the end of its useful life, a "normal replacement" action. The program may also motivate the customer to use more efficient new equipment when production capacity is increased or when new controls are added, e.g., EMCS, to existing equipment, a "new equipment" action. Free ridership can only occur when customers undertake "normal replacement" or "new equipment" actions. By definition, "early replacement" actions are those that the customer had no plans to undertake, eliminating any possibility of free-ridership.

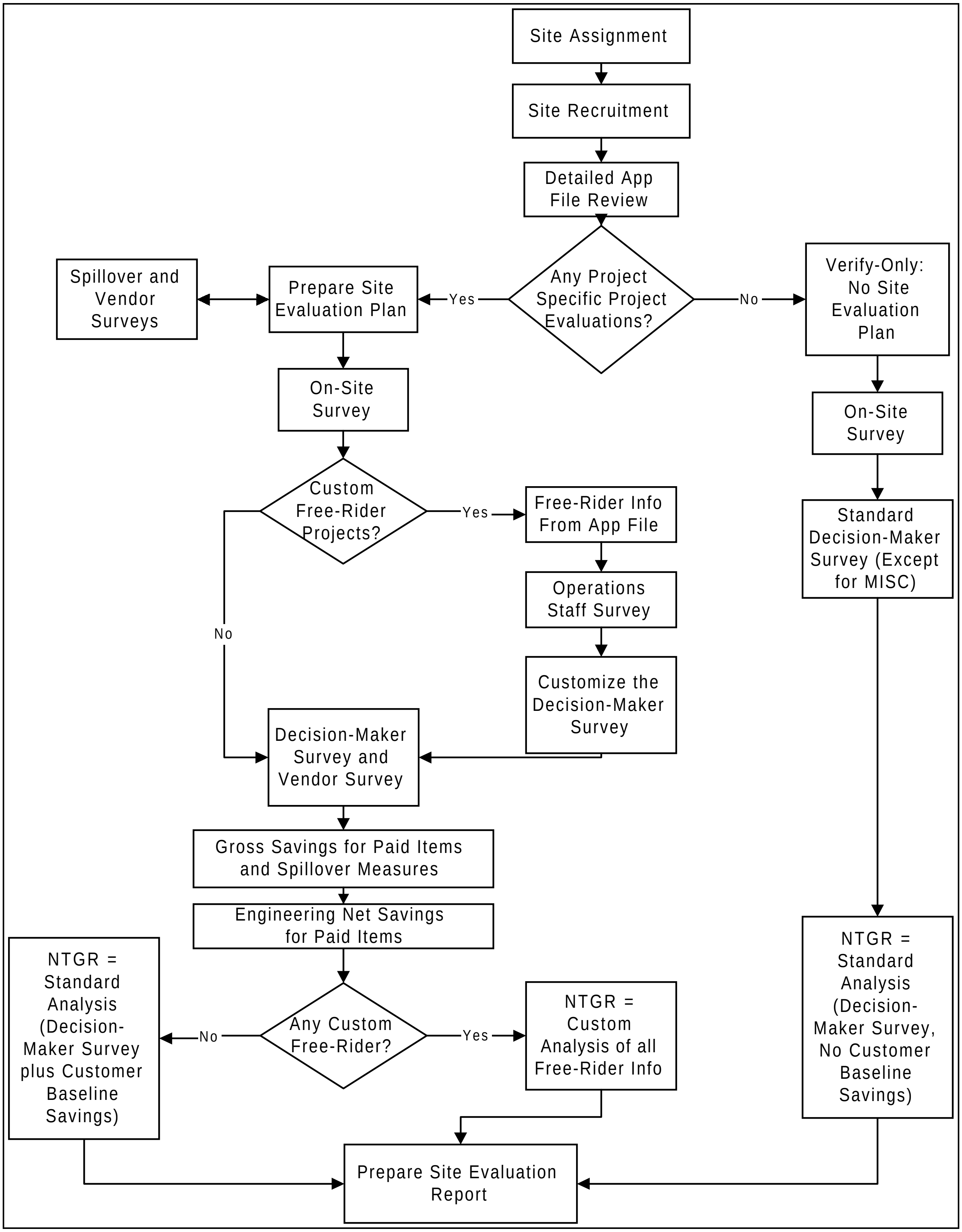
Full free-ridership, occurs when the participant would have implemented the **same** measure even if the program had not existed. Partial free-ridership occurs when the customer would have installed equipment which was more efficient than the equipment it replaces (or which is required by Title 24/20), without a rebate. However, that equipment would not have been as efficient as the equipment installed with the rebate. The equipment that they would have installed without the rebate constitutes the **customer baseline**. Full free-ridership occurs when the customer baseline is the same as the as-built conditions of the measure observed in the on-site survey.

Sources of information on Free-ridership

There are five sources of free-ridership information in this study.

1. **Program Files**. PG&E's program maintains a paper file for each paid application. These may contain information relevant to the analysis of free-ridership, such as information on the measure payback, or correspondence concerning the customer's motivations to adopt the efficiency measures.
2. **Operations Staff Survey**. Sometimes the person at the site most familiar with the operation of the equipment affected by a measure was not the decision-maker. In these cases, the Operations Staff Survey was administered to obtain an initial description of what the customer would have installed, if anything, in the absence of the rebate and the motivations for this action. This was used in forming customized free-ridership questions (CPS projects only), to be included in the decision-maker survey.

Figure 1: Summary of Site Evaluation Process



3. **Decision-Maker Survey**. When a site was recruited we also determined who was involved in the decision-making process which led to the installation of measures under the 1995 program. The Decision-Maker Survey obtained highly structured responses concerning the probability that the customer would have installed the same measure in the absence of the program. In addition, the survey obtained a description of what the customer would have done in the absence of the program, beginning with whether the installation was an early replacement action. If it was not, the decision maker was asked to provide a description of what equipment would have been installed in the absence of the program, which was used to define the customer baseline for the engineering calculation of net savings. If the decision-maker could not be sufficiently specific about the customer baseline equipment, we sought clarification from the person who participated earlier in the Operations Staff survey. The decision-maker was also asked to explain the customer's motivations for installing the efficiency measure. Additional questions were added to this survey for **custom project-specific** sites. These questions were based on information obtained either from program files or the Operations Staff survey, and were designed to confirm, clarify, supplement, or reconcile differences in the information obtained from these sources and to provide a deeper understanding of the decision making process.
4. **On-Site Survey**. During the On-Site survey our engineers observed the as-built and as-operated characteristics of the measures and the systems affected by the measures. With this information we modeled the efficient case energy use of the affected systems. Information from operations staff and the program file, along with applicable Title 24/20 standards, allowed us to model the pre-condition baseline. Information from the Decision-Maker survey allowed us to model the customer baseline, i.e., what they would have installed in the absence of the program.
5. **Vendor Survey**. In some instances, vendors were contacted following the decision-maker survey or the spillover survey. Some customers are not always aware that they are implementing energy saving measures that they otherwise would not have done in the absence of the program. This is most obviously the case when there is a participant who is not aware that vendors' recommendations have been affected by a DSM program. In this situation, the customer is not able to reliably self-report the influence of the program. The Vendor Survey was used to obtain information concerning the program's influence through the vendors that recommended or installed the equipment comprising the measure. The survey was completed for those respondents who indicate that PG&E exerted little influence (i.e., NTGR =< .3) on their decisions to install the items for which they received rebates.

Table 1 shows the data sources used in each of the three levels of free-ridership analysis. Although more than one level of analysis may share the same source, the amount of information that is utilized in the analysis may vary. For example, all three levels of analysis obtain data from the Decision-Maker interview. However, in the case of the custom project-specific analysis, the Decision-Maker interview contains additional site-specific questions that were used to clarify, confirm or reconcile information from other sources. The *standard* Decision-Maker survey is used for the other two levels of analysis.

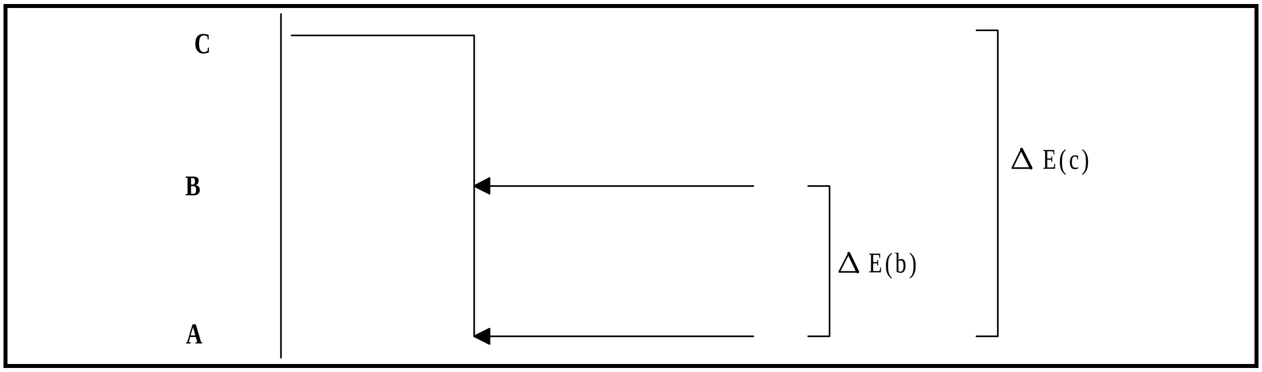
Analysis of Net-To-Gross Ratio (NTGR)

The net-to-gross ration (NTGR) is a dimensionless quantity that when multiplied by the estimate of gross savings for a project, yields an estimate of the projects savings net of the free-ridership effect. For CPS and SPS project the NTGR was first calculated using responses from the Decision-Maker Survey. This "self-report" method, is fairly common in studies where a comparison group is not available. Next, except in cases of early replacement where the self-report NTGR is set to 1, the ratio is adjusted based on an engineering evaluation of customer baseline savings, i.e., savings the customer would have achieved, if any, in the absence of the rebate.

Table 1: Information Sources for Three Analysis Levels

	Sources of Information					
Level of Free-Ridership Analysis	**Program File**	**Operations Staff Survey**	**Decision-Maker Survey**	**On-Site Survey**	**Vendor Survey**	**Spillover Survey**
Standard Project-Specific	■		■	■	■	■
Customized Project-Specific	■	■	■	■	■	■
Verification			■		■	

Figure 2: Three Basic Customer Options



What inaccuracies arise when the *unadjusted* self-report NTGR is used? Figure 2 presents the three basic options for any program participant. "C" refers to the old equipment, "B" refers to customer baseline equipment (some intermediate level of efficiency), and "A" refers to the efficient equipment that the customer installed through the program. "X" refers to the difference in usage between "C" and "B" while "Y" refers to the difference in usage between "B" and "A". The gross savings are defined as X + Y.

In the absence of the rebate, some customers face a normal replacement situation in which they must replace their old equipment and are considering equipment of varying efficiencies. If, in the absence of the rebate, a customer would have installed equipment less efficient than what it installed through the program, the greatest kWh savings that PG&E can legitimately claim is Y. Thus, multiplying the gross savings ("X" + "Y") by the self-report NTGR will *overestimate* the net savings since the rebate caused the customer only to go beyond the equipment with intermediate efficiency. An adjustment must be made in the self-report NTGR that reflects the fact that equipment of intermediate efficiency would have been installed in the absence of the program.

The calculation of an a customer baseline adjustment for normal replacement measures was conducted as follows: First, a few definitions:

E(a) = Energy use of as-built equipment.
E(b) = Energy use of alternative equipment, if considered by customer
E(c) = Energy use of pre-retrofit equipment

Next, various energy savings can be defined using these terms.

ΔE(a) = E(a) - E(a). This represents the savings experienced by a customer who would have installed the *same efficient equipment* in the absence of the rebate. Such free-ridership reduces the net savings to zero.

ΔE(b) = E(b) - E(a). This represents the gross savings assuming the baseline is the *alternate equipment,* i.e., what they would have installed in the absence of the rebate.

ΔE(c) = E(c) - E(a). This represents the savings experienced by a customer who would have kept its old equipment in the absence of the program. This is the gross savings assuming that the baseline is the *pre-retrofit equipment*.

Each of these three options has a probability of selection and they are defined as:

P(a) = probability of customer selecting as-built equipment without the rebate.

P(b) = probability of customer selecting alternative equipment without the rebate.

P(c) = probability of customer keeping pre-retrofit equipment without the rebate.

Using the above information, an adjusted self-report NTGR (ASR_NTGR) can be calculated under the simplifying assumption that the as-built equipment, the alternative equipment, and the pre-retrofit equipment are the *only* alternatives. Under this assumption:

$$P(a) + P(b) + P(c) = 1$$

The ASR_NTGR can then be calculated as:

$$ASR\_NTGR = P(a))[\Delta E(a)/\Delta E(c)] + P(b)[\Delta E(b)/\Delta E(c)] + P(c)[\Delta E(c)/\Delta E(c)]$$

This reduces to the following:

$$ASR\_NTGR = P(a)[0/\Delta E(c)] + P(b)[\Delta E(b)/\Delta E(c)] + P(c)[1]$$

or

$$ASR\_NTGR = P(b)[\Delta E(b)/\Delta E(c)] + P(c)$$

What is needed now are estimates of P(a), P(b), P(c), and ΔE(b)/ΔE(c). The estimate of P(a) is derived as 1 minus the self-report NTGR. At this point, it should be emphasized that we recognize the superior reliability and validity of the information contributing to the calculation of P(a) over P(b) and P(c) since it is based on the self-report NTGR derived from the Decision-Maker Survey. One of the questions on the decision-maker survey asks customers whether they considered any other alternatives to the equipment that they installed through the program. The options are basically two: 1) replacing old equipment with equipment that was not as efficient as the equipment installed through the program, and 2) keeping the old, pre-retrofit equipment. If option #1 is chosen, then P(b) is derived as 1 - P(a). If option #2 is chosen, then P(c) is derived as 1 - P(a). The advantage of this approach is that the information contributing to P(a) is allowed to drive the calculation of P(b) or P(c) in all situations. Of course, $\Delta E(b)/\Delta E(c)$ is based on engineering information obtained from customers regarding their old pre-retrofit equipment and what equipment they would have installed (customer baseline), if any, in the absence of the rebate.

Site-Level Spillover Assessment

Spillover is defined as the gross savings of measures that are not counted as part of the gross program savings and yet would not have been installed if the PG&E DSM programs had not existed. Spillover was only evaluated at those sites where a project-specific evaluation was performed. During development of each site evaluation plan, the customer was interviewed to determine if spillover was present. The objective of this spillover interview was to determine if other efficiency measures were implemented in 1995 due to the influence of PG&E programs, but were not rebated under the programs.

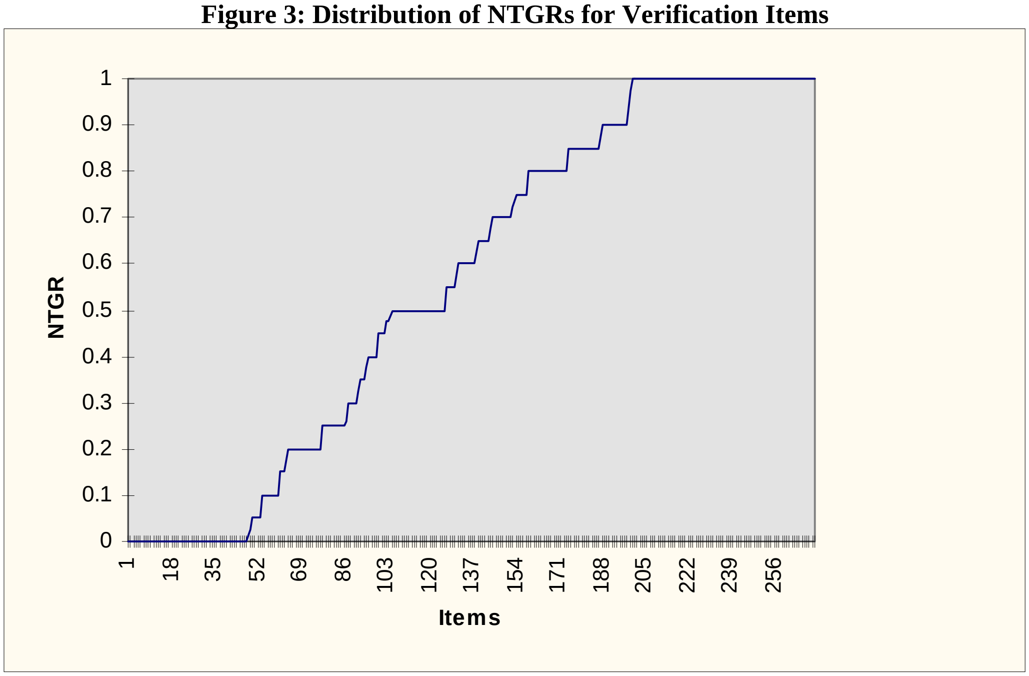
In order to measure the kWh savings associated with the spillover item, an on-site survey was required. However, given time and budget constraints, in order for a piece of equipment to be examined during the on-site as an instance of spillover, we had to be at least 60% certain that PG&E caused this installation to occur. This probability was provided by the respondent's answer to a question on the spillover survey. On the other hand, if this probability was less than .6, it might have meant that the respondent was unaware of the role that PG&E's influence on the vendor may have played in the selection of the efficient equipment. That is, customers are not always aware that they are implementing energy-saving measures that they otherwise would not have done. This is most obviously the case when there is a participant who is not aware that vendors' suggestions have been affected by a DSM program. In this situation, the customer is not able to reliably self-report the incidence of spillover for non-program installations of efficient equipment. For those participant respondents who indicated that there was less than a .6 probability that PG&E caused them to install their *non-rebated* installations, we asked them the name the person who recommended the measure (a distributor, selector, or installer) and proceeded to contact this vendor.

For spillover measures with probability greater than .6 and for which savings were likely to be large, we developed engineering savings algorithms using methodologies similar to those used for calculating the savings of program measures. We inspected the spillover measures and collected the necessary data during the on-site survey and developed an engineering estimate of spillover savings using the defined algorithms.

Results

For the 273 verification items the Self-Report NTGR was calculated based *only* on the responses to the decision-maker survey. The unweighted, overall self-report NTGR for these items was .56 with a standard deviation of .38. Figure 3 presents the distribution of the NTGRs.

Figure 3: Distribution of NTGRs for Verification Items



Standard Project-Specific NTGR Results

For 390 project-specific items, the self-report NTGR was first calculated and then adjusted for partial free-ridership. The adjusted self-report NTGR was based on the responses to the decision-maker survey, plus information collected, when necessary, from vendors, and baseline information collected on site that was used to calculate the customer baseline level of consumption. The unweighted, NTGR for these items was .77 with a standard deviation of .28. Figure 4 presents the distribution of the NTGR for the 390 items.

An important question is whether the adjustment for partial free-ridership was significant or not. For 32 of the 390 project-specific items, decision makers indicated that in the absence of the PG&E Program, it was more likely that they would have installed an alternative piece of equipment. However, for only 7 (1.7%) of these items at three sites was it possible to calculate the customer baseline and therefore the engineering ratio necessary to adjust for partial free-ridership. Recall that a large ratio means little adjustment and a small ratio means a large adjustment. These ratios ranged from .07 to .93, with 5 of the 7 greater than .65. For the two cases in which the adjustment was large, the gross savings were small. Thus, the adjustment for partial free-ridership was trivial.

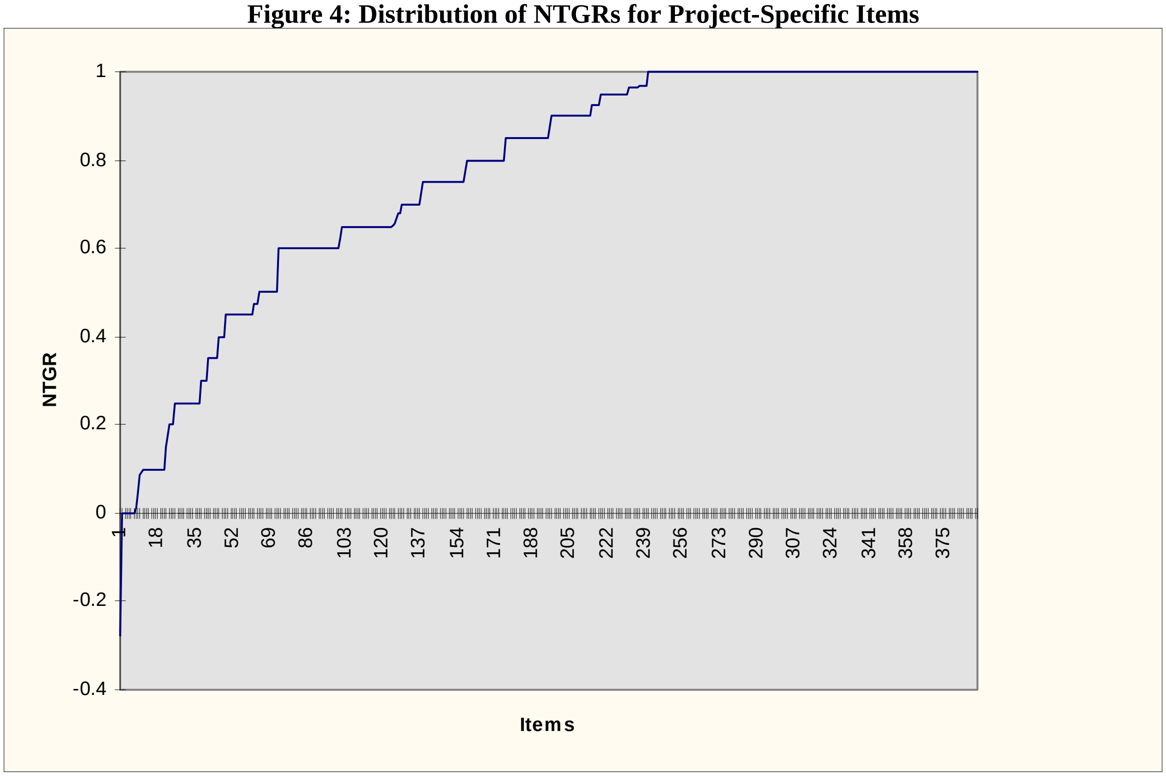
Note that it was not that respondents could not remember whether they considered alternatives; only 24 gave "don't know" as their answer. The overwhelming majority, 80%, specifically stated they did not consider any alternatives. If these program participants are unique in this respect, then future studies may be able to measure more partial free-ridership. However, if they are not unique in this respect, then partial free-ridership may be a relatively minor issue.

Custom NTGR Results

The primary purpose of the custom analysis was to seek additional information for the larger sites so that a more complete picture of the conditions surrounding the installation of the efficient equipment could be gained. This additional information could then be used to *modify* the Standard NTGR or *strengthen* the Standard NTGR. For the 146 custom items for which data were available, the unweighted Standard NTGR was .735 with a standard deviation of .25.

In the custom analysis of the 146 items examined, the Standard NTGR was modified for 42 items. Of these modifications, 35 were increases and 7 were decreases. Across all items, the changes produced a small *increase* of .019 in the overall, raw, unweighted Standard NTGR thus yielding a Custom NTGR of .754. Approximately half of this increase is due to one site, #402. At this site, the NTGR changed from -.26 to .95 (a change of 1.21) for reasons having to do more with the baseline than with our estimate of PG&E's influence. If one ignores this case, this increase of .019 is reduced to .009 producing a raw, unweighted NTGR of .744. For the remaining 104 items, the Standard NTGR did not change since any information identified provided insufficient grounds for *changing* the Standard NTGR or served only to *confirm* the Standard NTGR.

Figure 4: Distribution of NTGRs for Project-Specific Items



Early Replacement

All decision makers associated with the 672 items were asked a question concerning the timing of the installation. We were attempting to determine whether PG&E, in effect, caused the customer to accelerate the installation of the equipment. In cases where the installation was accelerated (early replacement) by at least one year, the NTGR was set to 1. Of the 672 items, 130 (19.3%) were early replacement.

Conclusions

A number of previous studies of non-residential energy efficiency programs have been performed using a self-report methodology to evaluate net impact. Typically the self-report information is used to form a net-to-gross ratio that accounts for free-ridership. In this study we added a number of features to this common methodology. This included an assessment of spillover savings, a customer baseline adjustment for self-report NTGR and customized probing to determine the details of the customer decision-making process for projects that resulted in large gross savings. These three features resulted in small modifications to the basic self-report NTGR. In part this was due the introduction of compensating adjustment, i.e., upward adjustment to the NTGR for some cases canceling the affect of downward adjustments in other cases. It is also due to the adjustments being small for most projects.

References

Pacific Gas and Electric Company, *Final Report: 1995 Nonresidential Retrofit Program: Industrial Sector Study*, March, 1995.