Evaluating Attribution, Causality, NEBs, and Cost Effectiveness in Multifamily Programs: Enhanced Techniques

Jody Myers and Lisa A. Skumatz, Skumatz Economic Research Associates, Inc.

ABSTRACT

As part of several evaluation projects, the authors have developed a number of useful steps to improve the attribution of impacts to program interventions. This paper presents an early application of enhanced evaluation methods to the multi-family sector – illustrating attribution/causality work for a multifamily new construction program. The methods and results are presented in this paper, and include:

- Theory-based evaluation, indicators, and hypothesis testing;
- Estimation of gross impacts and attribution of impacts using distribution and ranges of measure and intervention impacts, rather than less reliable point estimates. This uses more of the information gathered during data collection, provides a band of confidence around results, and more fully reflects the range of impacts induced by the programs.
- Enhanced measurement of naturally occurring adoption, multiple types of spillover/indirect market effects, and other impacts.
- Examination of the results across program types to explore patterns and important differences.
- Cost effectiveness assessment.

These extra steps improved the reliability and robustness of the results of the causality analysis and provided a better foundation to guide benefit-cost analysis and program and investment decisions – an important goal of an evaluation. The paper highlights benefits and impact of these approaches. We use the data to estimate the total net program impacts including net energy impact and net non-energy benefits (NNEB) for the program – information that was used in estimating the program's cost-effectiveness.

Introduction and Background

As part of work to conduct evaluations of several multifamily evaluation projects, the authors developed a number of analysis steps to help enhance evaluation results and applications. The results adapted procedures from residential and commercial applications to the separate case of multifamily dwellings – and in one case, a program with complexities embedded in the definition of "baseline." We adapted and applied robust approaches to examine free riders/naturally occurring adoption, spillover/market effects, and attribution – relying not just on point estimates, but specifically incorporating distributions and uncertainties to provide a better foundation for guiding program-related decision-making. This paper summarizes suggestions for improved methods for conducting attribution, non-energy benefits (NEB), and evaluation research in the multifamily sector.

Providing logical data/information that illustrates a consistency with causality can help formulate persuasive arguments demonstrating program effects, despite the fact that the literature takes pains to note that causality cannot really be proven. Key logical steps include:

- Assessing whether the technology works in the field;
- Using reliable methods to estimate associated effects and demonstrate causality, employing multiple methods to triangulate and confirm results if possible; and
- Reporting bounds and confidence intervals to improve credibility and useful application of the results.

Point estimates are less helpful than ranges or confidence intervals, which provide more robust and reliable information about impacts and can be used with greater confidence as guides for decision-making about programs, budgets, impacts and other decisions.

Steps for Attributing Program Savings

Of the two key methods of providing information on net to gross results – statistically-based difference of differences approach and self-reports from surveys – this paper focuses on the latter approach.

First Steps – Organizing Through Definitions, Objectives, and Outcomes

The attribution analysis is often part of a larger program evaluation, potentially including process evaluation, market characterization or attribution, and program attribution/impact analyses. We find two key steps are important to providing a well-designed, highly defensible analysis of attributable program savings.

- **Definition of participants, non-participants, and actors**. We take a systematic approach to defining participants and non-participants. Although these definitions may appear obvious, in many cases they are not. We construct a matrix that forces the analyst to clearly define participants and non-participants for each actor type related to spillover and free ridership (FR). This process helps further identify survey groups and commonly leads to identification of "partial" participants that need special treatment in the analysis. To provide the most robust analysis of attribution, we find it important to interview not just "owners," but other decision-makers that have a role in equipment specification or design.[1] This provides an opportunity to confirm results, obtain ranges and bounds, and explore the implications of differences in results by actor type.[2]

[1] Thus, for multifamily projects we commonly interview samples of building owners, A&E/builders/developers/contractors, staff, and where possible, tenants. Each "actor" group provides estimates of elements of net-to-gross (NTG), and results for each group can be weighted to best represent NTG. The results can be used to help address the "bounds" or range for the results for free ridership, spillover, and NTG. Interviewing non-participants is important to providing high quality estimates of non-participant spillover; however, participants can also provide indicative information on this component of spillover.

[2] And, in particular, occupants can be important interviewees for information on non-energy benefits for multi-family measures.

- **Mapping of researchable issues, outcomes, indicators, and data sources**. The review of the program's logic can illustrate the key researchable issues to be examined in the overall evaluation. Mapping these to the potential data sources assures that the data collection efforts will be integrated and as cost-effective as possible. A matrix of logic-related outcomes, assigned indicators of progress, and the measurement method and source for that work helps assure the attribution work covers all key anticipated program effects.

In order to assess the impacts of programs in the multifamily sector, it is key that interviews be conducted with multiple actors. Decisions about equipment and program attribution may have rested in several hands – developers, A&E/contractors, builders, landlords, and tenants – before, during, and after the program.

These organizational steps are crucial to making sure large-scale evaluations are as integrated as possible and helps clarify the priority research issues that will and will not be investigated.[3]

Next Steps – Net-to-Gross Computations

This step involves quantifying the amount of the program's performance that is attributable to the program. In order to accomplish this, the gross savings estimates developed through the impact analysis work need to be adjusted by the net-to-gross (NTG) ratio. This ratio is used to provide appropriate adjustment for the program's net effect – specifically, to estimate the impact of the program *above and beyond what would have happened without the program*. Translating gross program-tracked energy and demand savings into just that share that can be specifically attributed to the program is a complex problem in evaluation. Net program impacts reflect gross changes adjusted to account for the combination of two main effects.

- **Net Effect**: A reduction in the gross effect for "free ridership," or that share of program participants that would have undertaken the efficiency behavior or implemented the efficiency measure(s) even without the influence of the program or its market interventions.
- **Market Effect**: An additive adjustment to gross impacts, accounting for the positive impact and increases in efficiency that the program may have on market actors and actions above and beyond direct program participants. Resource acquisition programs tend to call this spillover; market transformation literature adjusts the terminology to focus on induced or indirect market effects (Sebold et.al. 2001).

Virtually all evaluations attempt to determine effects above and beyond what would have happened without the program, or the changes due to the program. For this reason, free ridership (or net effects) is a key component. **Free ridership** addresses the set of program participants that would have purchased the energy efficient measure, or adopted the behavior, even without the influence of the program – that is, the program was not instrumental in the participant purchasing

[3] For example for one multi-family program, to address the nuances needed for the regulators, it was necessary to measure NTG results in relation to two separate baselines – standard building practices, and energy code. Up front organization assured that all data collection addressed these differences as needed.

the energy efficient measure. Given that the smallest this factor can be is zero, this factor always reduces the gross savings attributable to the program.

On the opposite end of the spectrum is spillover or market effect. This figure attempts to measure the impacts that the program caused in the market through the influence of the program. There are at least three kinds of spillover:[4]

- **Inside Spillover or Inside Project Induced/Market Effects**: Inside project spillover is defined as additional measures installed or practices used by the household that were not incented directly by the program but were installed because of the influence of the program. These are measures not included in the program records or accounted for in program savings computations.
- **Outside Spillover or Outside Project Indirect/Market Effects**: Outside project spillover is defined as additional efficiency equipment installed by participating owners in their homes or at other buildings – but was not incentivized through the program. This factor accounts for the increase in efficient measures/practices adopted because of the influence of the program even without direct incentives.
- **Non-Participant Spillover or Non-Participant Market Effects**: Non-participant spillover refers to non-program energy efficiency (EE) measures purchased by non-participants that were inspired to purchase the measures because of program advertising or because more energy efficient equipment is in the market due to program actions.

By combining these measures, the net-to-gross computations can then be completed. Net program impacts are computed by applying adjustment factors for the effect of free riders and spillover or free drivers to the gross savings estimates for the program. The basic equation for the Net-to-Gross (NTG) ratio is: *NTG ratio = (Net Factor) x (Market Effects Factor).* The net factor equals the attributed fraction of savings, or the value one minus those savings deemed to be free riders, and the formula is *Net Factor = [1 – (free ridership)]* [5].

The market effects factor is a combination of program spillover factors that may influence actions taken outside of the program. The market effects factor is the sum of one and the spillover components: *Market Effects Factor = [1 + (inside project spillover+ behavioral changes) +outside project spillover + non-participant spillover]*.

Estimating Free Ridership Using Self-Report Survey Data

Given that free ridership figures are necessary for accurate net-to-gross calculations, our research has dedicated a number of specific survey questions to providing related information. The two basic approaches were to 1) ask how much of the savings or measures would have been installed without the program, and 2) the likelihood that measures of the same or better efficiency would have been installed without the program. Both of these answers provide key

[4] There can also be "Other Market" spillover/indirect effects which can occur through several pathways. For example, manufacturers may change the efficiency of their products, and/or retailers and wholesalers may change the composition of their inventories to reflect the demand for more efficient goods created through an energy efficiency program. Another example might be new building codes or appliance standards adopted in part due to the demonstration of technologies through an energy efficiency program.

[5] Note that some researchers and regulators add the two effects rather than multiplying them in computing NTG.

information related to decision-making and free ridership, and each is well-suited to certain situations.

Our research has successfully approached this task by asking about the proportion of the program measures that would have been installed without the program. This provides the most direct response to the concept of free-ridership. However, in some cases, the measures are "indivisible" – that is, a boiler may be either installed or not, but a portion of a boiler project is not installed. In other cases, the project may be a "go/no-go," and the program was the factor that pushed the entire project to "go." When asked share of measures in this case, the respondents are tempted to respond with "all/100%" or "none/0%." Asking both questions provides us valid responses for both types of situations. Further enhancements include asking respondents:

- For a high and low range as well as their best estimate of the share of savings that would have occurred without the program's influence. This helps walk them through the logic of their response;
- The likelihood that the owner would have installed measures of the same or higher efficiency without the program;[6]
- The share of measures (and resulting energy savings) they would have installed without the program – asking a range and best estimate; and
- The share of energy savings they installed because of the program.

Using a combination of the questions used to assess free ridership has allowed us to derive the base value by "filling in" missing data with a hierarchy of responses. This is especially important since not all respondents will be able to answer all questions. This approach can be used to compute the "base" free ridership estimate.

Using a variety of survey methods can help confirm the validity or consistency of responses provided to questions about free ridership and spillover. For example, some argue that over time, participants may have convinced themselves the savings would have been installed without the program. Asking related questions can allow examination of the consistency of those self-reports on free ridership. In multifamily research, we have included the following types of related/influential questions to further examine free ridership estimates.

- **Importance of Incentives**: Respondents were asked the importance of the program's financial (or technical) assistance in their decision to install the energy efficient measures. If respondents say the assistance was very important, but also provide a high free-ridership value, the results are inconsistent and the free ridership value may need to be adjusted downward.
- **Prior Plans to Install**: Respondents were asked how firm their plans were to install the same efficiency of measures before they heard of the program. If plans did not exist (or were not very firm) or were for lower efficiency measures, but they report high free ridership, the FR value may need adjustment downward.

[6] This version of the question is more appropriate when the measure(s) come in large increments. A boiler is installed or not installed – therefore, respondents would be likely to answer 0% or 100% to the question about percent of savings. Percent likelihood that 100% of the savings would have been installed provides better distribution of responses with better estimation properties.

- **Increases in Efficiency**: The respondents were asked if the program increased the efficiency of the measures that were installed beyond what would have been installed without the program. A report of a program-induced efficiency increase may be inconsistent with high free ridership.
- **Timing of Investment**: If the program caused efficient measures to be installed considerably earlier than would otherwise have occurred, then there is no free ridership for at least the period that the measures were moved forward in time.[7] If the installation occurs considerably earlier, it might also be argued that uncertainties might have jeopardized their installation, and high free ridership values may need to be adjusted downward.

Taking into account influencing factors through systematic examination of individual responses can improve the confidence in the estimate of free ridership and provide a more robust range for the value.

Estimating Spillover Using Self Report Survey Data

As mentioned above, survey responses can be used to compute three types of spillover or indirect market effects. A three-step approach was employed to develop the estimate of each spillover component.

- **Existence**: Respondents are asked about the existence of the effect (inside, outside, or non-participant spillover). This is prior to asking about size, and provides at least partial information if they are not able to answer more quantitative questions.
- **Extra "Savings" Installed**: A battery of questions is asked that allows estimation of the level of extra spillover savings (as a multiple of the direct program-installed savings). The specific questions vary based on the spillover type.[8]
- **Confirming Share Due to Program**: Finally, as a check, we reconfirm what portion of the spillover factor they identified in the "extra savings installed" is actually attributable to the influence of the program.

Since these questions are asked of a number of actors – owners, occupants, developers, A&E, and others – these data can then be analyzed to determine an estimate of the range and best estimate of elements of spillover.

Computing non-participant spillover is more complex, although none of these computations is simple. Assigning non-participant[9] spillover to any specific program in a utility's portfolio is difficult or impossible because of the potential for overlapping and cross-over influences. Respondents – and even more so, non-participants – do not always clearly distinguish between the various utility programs or the influence of specific elements of the portfolio. Therefore, while they are asked about a specific program, the results must be used

[7] Since one of the issues associated with free ridership is accounting for savings, it is important to consider the timing of savings resulting because of the program.
[8] The questions for non-participant spillover are somewhat complex. For best results, we ask about non-participants about non-participant spillover, with information from participants as auxiliary data on this factor.
[9] And sometimes untangling outside spillover effects can also be difficult.

cautiously because of the overlap between programs within one "sector." Given that there are usually only a limited number of multifamily programs, direct overlap may not be a large problem. However, if other programs affect builder practices, or if other programs have the effect of changing the stocking practices for energy efficiency measures, overlap can still occur. Therefore, deriving reliable estimates of non-participant spillover is difficult and, although self-reports may be the best way of obtaining the information, the work may only be able to provide an indication of the magnitude of outside spillover.

Net-to-Gross (NTG) Ratio: Computation and Benchmarking

As indicated by the formulae presented earlier in the paper, the computation of the net-to-gross adjustment factor follows directly from the estimates of free ridership and spillover. Information by actor group, as well as overall ranges and confidence intervals can be computed.

It is often useful, however, to compare or "benchmark" the results against free ridership, spillover, and net-to-gross results from other programs in order to provide the strongest interpretation of the results. Skumatz Economic Research Associates, Inc. (SERA) has assembled a library of more than 100 results on NTG from programs across the U.S. and limited international results. We sort these results to identify "like" programs or measures and program designs in order to support comparison of the results with other programs. If the results differ, it is important for the researcher to explore the reasons for these differences. For example, differences can be expected for rebate vs. technical assistance programs, but if similarly designed programs derive very different results, the utility will want to know whether the differences are due to outreach or targeting variations, or because one program included more "mainstream" measures and there was less free-ridership from a program that focused on more cutting-edge measures. These results are important to a robust (and useful) analysis of program attribution.

These basic techniques have been adapted by the authors for a broad range of programs, including residential product outreach and product incentive programs; new homes and remodeling programs; commercial new construction; renewables, and other programs. Enhancements being examined that include scenario analysis and decision tree techniques to "bound" the effects, and options to use a Bayesian approach addressing "degree of belief" in the impacts. Thus, the attribution work is incorporating several methods of addressing risk – an important component of using causal results.

Example from Multifamily Programs Alternatives

The estimated Net-to-Gross Ratios for three template multifamily programs is provided for illustration purposes. Table 1 shows results from an Energy Star® Multifamily program, an advanced metering program, and a low income retrofit program.

The results can be compared across programs and to programs at other utilities or in other regions to better inform program conclusions; for clients, we generally compare to results from data we have collected on more than 100 programs across the nation. For example, the results here show relatively high free ridership for the Energy Star® program (when compared to Energy Star® programs elsewhere). This particular program had a baseline loophole that required few developers to change the measures installed over standard practice; when that loophole is adjusted, the program might be expected to show free ridership values closer to the 0.2 found in other locations. Free ridership for the other programs is lower, and show sizes that make sense given the programs.

Table 1. Summary of NTG Elements and Computation of NTG Ratio

		Spillover or Indirect Effects			(1+Col B+C+D)	(1-A)*(E)
	A. Free Rider (rounded)	**B.** Inside	**C.** Outside	**D.** Non-Partic.	**E.** Market Factor	**F.** NTG Ratio
Energy Star® Program	0.50 (0.4-0.6)	0	0.1-0.2	0-0.15	1.2 (1.1-1.4)	0.6 (0.4-0.8)
Advanced MF Metering Program	0.1 (0-0.35)	0.1-0.3	0.1-0.2	0-0.1	1.4 (1.2-1.6)	1.3 (0.8-1.6)
MF Low Income Retrofit	0.3 (0.2-.4)	0	0-0.1	0-0.1	1.1 (1.0-1.2)	0.8 (0.6-1.3)

Advanced meters are rarely installed without program incentives and low income buildings tend not to buy new measures until it is urgent (low free ridership). Spillover values were relatively low; many retrofit programs in the literature show values of 0-0.5. The resulting NTG ratios are presented, providing estimates of the program's attributable effects.

By comparing with other programs, we can examine whether differences can be explained, whether results are reasonable, and possibly tease out lessons for improving the program. For example, in some cases overlapping programs within a sector might affect the free ridership; other causes may lie in the level of how "cutting edge" is the equipment incented under the program.

Non-Energy Benefits (NEBs): Enhancing the Robustness of the Evaluation

Non-energy benefits are increasingly being studied as their worth has been more adequately recognized and measured, despite the more direct nature of energy savings[10] and other metrics as indicators of program effects. These "omitted program effects" include any and all impacts that are not directly the energy and bill savings resulting from the program. Previous work shows that these benefits are significant in relation to the energy savings, and are highly valued by participants. In some cases, the analysis suggests that the primary value from the program was non-energy benefits, rather than energy-related bill savings. Previous work also indicates that market actors – specifiers like builders, architects, engineers, and contractors – also recognize these benefits and use them in "selling" energy efficiency.

Although the literature calls them non-energy benefits, they include the "net" of both positive and negative effects that may be attributable to the program. The convention has been established to separate these benefits into three "perspectives" (Skumatz & Dickerson 1997):

- **Utility NEBs**: These include utility/ratepayer-type benefits result in reduced revenue requirements, including savings in a variety of administrative and carrying costs related to arrearages, service terminations, and related changes, as well as reductions in T&D losses when fewer kWh are distributed through the system. The changes attributable to these impacts are mostly valued at utility avoided costs for the relevant labor category, etc.
- **Societal NEBs**: Societal benefits include the value of reductions in emissions, economic stimulus, and similar public benefits. The values associated with these program-caused changes vary with the type of impact.

[10] And, depending on the scope of the evaluation, awareness, market share, and other metrics.

- **Participant NEBs**: Participant impacts include effects above and beyond energy savings, and include improvements in comfort, lighting quality, resident satisfaction, equipment maintenance benefits, safety issues, and a wide variety of other NEBs. While many of these indirect benefits may be difficult to measure, they can ultimately be translated into dollar terms, and incorporated as net program benefits accruing to participants.

There is a substantial literature addressing utility and societal NEBs, and estimates of these impacts can be computed from a combination of primary and secondary data (Skumatz & Dickerson 1997). Assessment of the participant portion of the benefits is the most challenging portion of non-energy benefits work. The authors have spent considerable time on this issue, and based on extensive review of measurement literature and field testing of approaches, have developed a series of highly defensible contingent valuation, scaling, discrete choice, and other revealed and stated preference approaches to estimating these "hard to measure" (HTM) impacts (Skumatz & Gardner 2006).

While these measurement methods can be complex to implement, the work significantly enhances the usefulness of the evaluation efforts, as demonstrated by the authors' work on applying these techniques to more than 50 residential, multifamily, commercial/industrial, renewables, and demand response programs across the U.S. and internationally. However, in addition to measurement issues, providing reliable estimates of the NEBs actually attributable to a program requires that the results must be "net" in several key ways described below (nuances that have not been incorporated in other NEB research).

- **Net Positive and Negative**: Despite the historical name for these impacts (non-energy benefits), both positive and negative impacts must be incorporated.[11]
- **Compare Efficient to Standard Equipment**: To attribute the impact due to the program, the respondents need to be asked about the NEBs for the new efficient equipment relative to the base non-efficient equipment that would otherwise have been purchased. The appropriate comparison is generally not the new efficient equipment but the old equipment that was in place.[12]
- **Net of Freeriders**: Similarly, if there are free riders that would have purchased the same equipment without the program, then the NEBs associated with that equipment should not be attributed to the program.

NEBs Results from Example Multifamily Programs

Our research on programs across the nation and internationally demonstrates that, depending on the program design, the NEBs can be quite large – ranging from about 50-400% of the value of other energy benefits in the benefit-cost equation (Skumatz & Gardner 2005). The

[11] The term we use is "net non-energy benefits" (NNEBs) but we will refer to them as "NEBs" in this paper. Over a 10 year period, we have developed effective (proprietary) methods of asking these questions and valuing the responses. In addition, a model "NEB-It" © SERA 2001 is used to compute values.

[12] However, some caveats are needed, depending on how the work is to be used. It may be that in the case of residents that would not have purchased new equipment at all without the program, a case may be made that for participant NNEBs, they recognize all the change from old equipment to the new efficient equipment. Also, if the measures would not have been installed for a period of time, the full NNEBs may be appropriately credited (as should the savings) during the interim. However, these are fine points on the principles discussed above.

benefits to participants alone can equal half to almost twice as much value as the direct energy savings.

Exclusion of NEBs from program evaluations and the decisions that derive from them has been commonplace in part because NEBs are difficult to measure. However, that practice leads to the implicit use of the value of "zero" as the NEB. The results from numerous studies show that this simplification can be misleading.

Certainly the results and currency valuations vary by type of program, measures included, weather, and program targets. For instance, we find programs targeted at households with chronically ill inhabitants to lead to higher benefits from measures that address comfort. We find including gas measures can increase the NEBs associated with the program. Comfort-related NEBs are higher in climates with higher heating (or cooling) needs.

The authors conducted a detailed non-energy benefits (NEBs) analyses on a number of MF programs; template results from three different types of programs are illustrated in the Table below. The examples include an Energy Star® Multifamily program, an advanced metering program, and a low income retrofit program. Table 2 summarizes an assessment of the direction and size of NEBs for the programs.[13]

Table 2. Direction and Proportion of NEBs' Impacts from Sample Multifamily Programs

	Energy Star®			MF Metering Program			Low Inc. Retrofit
Category	**Negative**	**Positive**	**% of Total NEBs**	**Negative**	**Positive**	**% of Total NEBs**	**% of Total NEBs**
Operating cost (excl. energy)	0%	69%	14%	10%	45%	6%	0%
Equip maintenance	0%	53%	10%	10%	35%	8%	-13%
Equip performance	0%	47%	9%	4%	32%	4%	14%
Equip lifetime	0%	62%	10%	5%	10%	2%	3%
Occupant satisfaction	0%	56%	9%	21%	50%	11%	1%
Occupant comfort	0%	44%	8%	33%	19%	4%	3%
Aesthetics/Appearance	0%	31%	5%	12%	12%	19%	17%
Lighting/Quality of light	13%	31%	5%	8%	8%	2%	13%
Noise	6%	44%	6%	4%	4%	0%	8%
Building safety	0%	12%	2%	0%	13%	3%	5%
Ease of leasing/selling	0%	47%	6%	14%	46%	10%	17%
Doing good for environment	0%	81%	12%	5%	76%	16%	27%
Power quality/reliability	0%	33%	4%	0%	70%	10%	
Labor requirements				11%	56%	11%	
Ability to stay in units/avoid moving due to bill consistency							5%
NEBs as multiple of energy savings			67%			44-110%	108%

[13] Total non-energy benefits estimates were obtained using a labeled magnitude scaling technique. Respondents were first asked whether they experienced negative, positive or no effects with respect to both individual categories of non-energy benefits and non-energy benefits in aggregate. They were then asked to value those benefits relative to the energy savings that they experienced. The work uses proprietary multipliers developed from data from thousands of surveys © SERA.

Environmental benefits, operating cost, equipment lifetime, occupant satisfaction and equipment maintenance were among the most positive categories associated with the programs. In each of these categories, high percentages of those surveyed reported positive effects. The effect of "doing good for the environment" was rated as positive by particularly large percentages of respondents, punctuating the idea that some of the most important consequences of the program are distinct from the goals of increased energy efficiency and cost savings, at least in the eyes of program participants. Few reported experiencing any effects, positive or negative, in building safety, aesthetics or power quality – except in the case where one of the program's objectives was to install meters that improved power quality. Programs can also learn from negative (and low valued) effects. Negative effects were reported for noise (6%) and light quality (13%) in the Energy Star® program, and occupant satisfaction and comfort for the metering program; however, for each of these categories a greater percentage of respondents reported experiencing positive effects than did negative effects.

Table 2 also shows the proportion of the total NEBs reported by program participants attributable to the same program effects categories.[14] After the environment, highest value NEBs included operating cost and equipment performance/maintenance issues, comfort-related benefits, and benefits to the selling/leasing value. Significant maintenance concerns were noted in one program (-13% of the value). The NEB analysis identifies areas of program strengths (with marketing and program design applications) and weaknesses (which may indicate a need for reconsideration of measures, or additional interventions to address problems). Finally, the evaluations show that the value of the NEBs to participants varied between about 40% to greater than the energy savings from the program, depending on the program type and focus.[15]

Summary and Implications

Attribution

The authors employed several useful steps to improve the robustness of the analysis of market effects attributable to measures throughout dozens of recent projects for different residential and commercial programs. The focus was on using theory-based evaluation, but also worked toward providing a burden of proof on par with the criteria for other public and private investments. Four steps were undertaken for this effort. We use an enhanced method to address induced market effects/free riders, spillover, and free drivers to allow for partial free riders and used indicator methods to provide evidence on program-induced effects on spillover and free drivers. We address the issue of uncertainty in the attribution work by interviewing various actors, asking for ranges, and using scenarios to "bound" the effects, providing results that are more useful for causal analysis and for supporting scenario analyses to assess risk. Finally, we also compare the results to programs elsewhere to identify anomalies and better assess the program's relative performance.

---

[14] It is important to note that other programs may give rise to non-energy benefits that do not fall into the categories used in our survey. Our category selection was based on pretests that explored the kinds of benefits experienced by participants. In other work we have identified the 5-8 underlying factors that drive the NEBs and argue to focus benefits in those areas to minimize overlap. We also specifically request interviewees to note overlap issues.

[15] These are the NEBs unadjusted by the net-to-gross ratio. Adjusted NEBs are easily computed applying the NTG ratio to this figure.

The approach outlined in this paper used significantly more data gathered and more fully reflects the range of impacts induced by the program. Although causality can never be proven, these extra steps improved the reliability and robustness of the results of the causality analysis and provided a better foundation to guide program and investment decisions – one of the most important goals of an evaluation.

Non-Energy Benefits

NEBs are an important set of benefits provided by energy-related measures and features in multi-family buildings, although they are, unfortunately, often ignored. These NEBs can be measured, using several reliable techniques, and they represent effects that are attributable to programs. Utilities may run energy conservation programs to reduce energy use, and builders may build multifamily dwellings that include energy saving features, measures, and designs. However, energy savings may not be – and appear not to be – the highest valued outcome of these measures and features to buyers/participants. NEB results have several applications.

- **Benefit/Cost Analysis**: Although benefit-cost analysis was the focus of much early NEB work, this is generally not the primary application. However, some share of these effects may be appropriate for inclusion in program benefit cost analysis or regulatory "tests." Some programs are assessing B/C scenarios that incorporate 10%, 25%, and 50% of estimated NEBs to provide a fuller assessment of their program's impacts.
- **Program Design**: NEBs can be incorporated into initial decision-making about which measures/features to include in new/remodeled buildings (or into programs) and computations of costs and benefits from investment in energy using equipment. They can also be used to identify and address key program barriers. The analysis of barriers (negative NEBs) is richer than traditional process evaluation results on barriers as they can provide sophisticated analysis of the dollar values of the interventions that may be needed to "get past" the barrier[16] (e.g., rebates, warranties, Skumatz & Gardner 2005).
- **Marketing/Outreach/Advertising to Attract Buyers to These Efficient Buildings or to Energy Conservation Programs That Incorporate These Measures**: The value of these other program-derived effects (high value NEBs) – may be stronger selling points for the measures than energy savings – and these benefits should be used as key sales messages in program outreach (Skumatz & Gardner 2005). Proctor & Gamble doesn't sell households Tide laundry detergent based on "buy this because it gives us greatest profits," and the implication of the NEB results has shown that selling efficiency programs on energy efficiency/conservation which is important to program design is a poor approach – these are often not the highest valued benefits participants derive from programs, and energy efficiency may not be the most appealing to advertise. The NEB results also indicate that it may not be the most important feature people want to buy.

[16] Dollar values of importance are a much more useful way of assessing barriers than a 1-5 "importance" scale.

References

Sebold et al. 1997. "A Framework for Planning and Assessing Publicly Funded Energy Efficiency," Study ID PG&E-SW040 (calmac). Prepared for PG&E, San Francisco, Calif.

Skumatz, Lisa A., and Chris Ann Dickerson. 1997. "Recognizing All Program Benefits: Estimating the Non-Energy Benefits of PG&E's Venture Partner Pilot Program (VPP)." In *Proceeding of the 1997 Energy Evaluation Conference*, Chicago, Ill.

Skumatz, Lisa A., and John Gardner. 2005. "Methods and Results for Measuring Non-Energy Benefits in the Commercial and Industrial Sectors." In *Proceedings of the American Council for an Energy-Efficient Economy Industrial Conference*, West Point, N.Y.

_____. 2006. "Differences in the Valuation of Non-Energy Benefits According to Measurement Methodology: Causes and Consequences." In *Proceedings of the Association for Energy Service Professionals NESP Conference,* San Diego, Calif.; AESP, Clearwater, Fla.