Methods and Results for Measuring Non-Energy Benefits in the Commercial and Industrial Sectors

Lisa A. Skumatz, Skumatz Economic Research Associates, Inc.
John Gardner, Skumatz Economic Research Associates, Inc.

ABSTRACT

Based on the results of several dozen projects completed over the last decade, the authors have developed and refined methods for measuring non-energy impacts – both positive and negative – from commercial and industrial energy efficiency programs. This research covers measures of non-energy benefits (NEBs) in three main categories: utility/agency benefits, societal benefits, and participant benefits.

Participant benefits are particularly hard to measure (HTM). The authors pioneered, developed, and tested several methods of measuring these benefits; specifically, these methods included approaches such as contingent valuation; ordered logit; value to owner; comparative or relative valuations; willingness to pay (WTP) and willingness to accept (WTA); and scaled valuations, modifications to the Labeled Magnitude Scale.

This paper works to establish and detail a set of best practices for measuring NEBs, participant benefits in particular. The paper then applies those measurements to decision-making and policy formation. First, it reviews the NEBs concept, then examines various techniques to measure them, provides an empirical comparison of results from four key techniques (WTP/WTA, contingent valuation, relative and scaled valuations), and finally discusses the application of NEBs valuations to program evaluation.

Introduction: NEBs as Omitted Effects

While energy savings, energy efficiency awareness, market shares and other metrics provide direct indicators of a program's effect, there exists a significant body of work that has developed around recognizing and measuring net non-energy benefits (NEBs). Net non-energy benefits include all impacts that do not directly result from the energy and bill savings due to the program. Previous work shows that these benefits are significant relative 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 – such as builders, architects, engineers, and contractors – recognize these benefits and use them to help sell energy efficiency.

Defining NEBs

NEBs include a variety of impacts that result from energy efficiency programs. Although the literature calls them non-energy benefits, they are the "net" of both positive and negative effects that may be attributable to the program. As developed in Skumatz (1997), the convention has been established to separate these benefits into three groups:

- **Utility NEBs:** These are utility/ratepayer-type benefits that result in reduced revenue requirements, and includes: (1) savings through a variety of administrative and carrying costs such as related to changes in arrearages, service terminations; and (2) reductions in

T&D losses when fewer kWh are distributed through the system. Changes attributable to these impacts are valued at utility avoided costs for the relevant labor category, etc.

- **Societal NEBs:** Societal benefits or public benefits include program results such as economic stimulus or the value of reductions in emissions. The values associated with these program-caused changes vary with the type of impact.
- **Participant NEBs:** These benefits extend beyond energy savings and include a wide range of non-energy effects, including: improvements in comfort; lighting quality; equipment maintenance; and many others. While many of these indirect benefits may be difficult to measure, this paper addresses how they can be translated into dollar terms and incorporated as net program benefits accruing to participants.

Table 1. Net Non-Energy Benefits (NEBs) Categories: Changes from Program Interventions

Utility Benefits	
Reduced carrying cost on arrearages (interest) Bad debt written off Shutoffs / Reconnects Notices Customer calls / bill or emergency-related Other bill collection costs	Emergency gas service calls (for gas flex connector and other programs) Insurance savings Transmission and distribution savings (usually distribution only) Power quality / reliability Reduced subsidy payments (low income) Other
Societal Benefits	
Water and wastewater infrastructure Economic benefits–direct/ indirect multipliers Health and safety equipment	Emissions / environmental (trading values and/or health / hazard benefits) Other
Participant Benefits[1]	
Commercial/Industrial Participants	
Water / wastewater bill savings Operating costs (non-energy)[2] Equipment maintenance Equipment performance (push air better, etc.) Equipment lifetime Productivity Tenant satisfaction / fewer tenant complaints Comfort Aesthetics / appearance Lighting / quality of light Noise	Safety Ease of selling / leasing Product losses (mostly refrigeration at grocery) Labor requirements Indoor air quality Health / lost days at work Doing good for environment Reliability of service / power quality Savings in other fuels or services (as relevant) NEGATIVES include: Production disruption during installation. Others are included above (some may have worse maintenance, etc.)

Typical categories of benefits based on past projects are presented in Table 1 above. This list is not comprehensive, and obviously some benefits can cross categories. In addition, in our research, we tend to not include tertiary type benefits such as tax–related impacts because we prefer to be more conservative. Whether specific benefits are included or excluded from a specific analysis depend on which measures are included in the program. The list of benefits used in the program attribution analysis is usually refined in collaboration with the program staff.

Note that benefits can arise in multiple categories without being redundant. The easiest example could arise on the commercial or industrial side, but is more common on the residential side. For instance, a reduction in bill-related calls to the utility helps both the utility / ratepayers *and* the participating businesses / households who make or receive those calls. This is not double-counting benefits – rather, it recognizes that some effects have multiple beneficiaries and each is valued at the appropriate tailored valuation method. Each benefits from less time on calls; the utility valued at customer service staff marginal labor rates; businesses at wage rates;

[1] Positive and negative impacts, estimated using participant surveys for many of the NEBs.
[2] Sometimes omit if likely to double count with the next two categories

the minimum wage rate for low-income households; and wages or leisure proxies for households. Benefits are recognized and realized by each group; the inclusion of the benefits in specific computations depends on the benefits' appropriateness to the application.

Attribution of utility and societal NEBs can be measured by using a combination of primary and secondary data. There is extensive literature that measures the arrearage impacts of programs (particularly low income programs), as well as many other impacts. Detailed examination of the program impacts – or the literature– may be needed to estimate the impacts on reconnections and other factors that may be affected by the program.[3]

Societal impacts also have a significant amount of literature and indeed, the two key components-- environmental and economic impacts – have a very high degree of volatility depending on the data sources and valuation methods used. Impacts on greenhouse gases (GHG) are increasing in importance and have been estimated in the literature. These impacts are a "slippery slope" – they can be estimated in a simplistic way, but if health impacts are to be measured in detail, then issues related to specific microclimates and time of day and zones are important. For some programs, average generation mix should be used to assess emissions; for others (e.g. a peak load reduction program, residential air conditioning programs, etc.) emissions from marginal peak load plants should be used to estimate changes in emission from the energy savings. Valuations are the source of considerable debate in the literature as well.[4] There exists literature that estimates economic impacts of energy efficiency programs. Some of the literature is flawed because they estimate the job creation and economic multipliers of a *gross* expenditure on the economy instead of measuring the net impact of a switch from, say, the sectors included in electricity generation, into the economic sectors affected by weatherization or other programs.[5]

Measuring Program-Attributable NEBs

To provide credible estimates of the NEBs actually attributable to the program, the results should represent "net" effects, measured in several specific ways.

- **Net positive and negative**: Despite the historical name for these impacts (non-energy benefits), both positive and negative impacts must be incorporated. We use NEBs to mean net NEBs. In each case, when the term NEBs is used in this paper, we mean net NEBs.[6] We generally ask questions in three levels – (1) whether they can name any benefits (to get at existence); (2) whether they experienced any impact from each of a set of NEB categories – positive, negative, or no impact; and (3) the size of the effects (positive or negative) using one or more of several approaches described in this paper.
- **Compare efficient to standard equipment**: To determine 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 otherwise would have been purchased. Appropriate comparisons are generally not to the old equipment in place.[7]

---

[3] See for example, Hall, Skumatz, and Megdal, 2000 provides an extensive discussion of these estimation methods.

[4] For some clients, there are values that have been agreed upon by the regulators. For others, we used specific values included in the literature, or averages of valuations from many sources. Which valuations are most appropriate depends on not only the location, but also the use to which the work will be applied.

[5] For an extensive discussion of the environmental and economic impacts, see Imbierowicz and Skumatz, 2004.

[6] We use NEBs (non-energy benefits) to mean net NEBs (NNEBs). In each case in this paper, when the term NEB is used, it is assumed to be "net".

[7] 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

- **Net of free riders**: 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. However, in parallel, a case may be made that the NEBs for the spillover purchases may be added to the NEB totals.

These nuances have not always been incorporated in NEB research and such refinements can have significant impacts on the results. It is not practical to ask detailed free ridership questions in all surveys, and in those cases, proxies can be derived. In research on net to gross ratios from commercial / industrial (and residential) programs across the nation, the authors identified common ranges for free ridership values based on different program types. Some of these results are provided in other publications by the authors (Skumatz et.al. 2005). Free ridership for some traditional high efficiency equipment rebate programs may be as high as 50%, which leads to a discount of the NEBs by about 50%. Values of 5% to 20% may be more accurate for newer programs that promote more cutting edge equipment or practices that lead to lower discounting of the NEBs. These adjustments are an important step in producing reliable and credible NEB estimates that are attributable to the programs.

The authors have conducted scores of projects to measure net NEBs for both residential and commercial / industrial programs – that include utility, societal, and participant benefits. The most challenging portion of non-energy benefits work is assessing the participant portion of the benefits. The authors have spent considerable time working on this issue, and have developed several credible methods of estimating these "hard to measure" (HTM) impacts. The topic is addressed more fully in the following section.

Approaches to Estimating Participant Non-Energy Benefits for Programs

Methods for measuring specific categories of participant NEBs depend on the category. For instance, NEBs related to water savings requires information on water usage impacts (using before and after data, impact evaluation, or deemed water savings per measure installed). These data are then valued using local water rates for the customer class for participant impacts.[8]

However, for the vast majority of participant impacts, participant surveys are needed. A few commercial/industrial program participants have sometimes conducted direct measurements of impacts; for example some have measured the impacts of lower maintenance / labor from new measures. However, these measurements have not been conducted for the vast majority of participants in programs, and using the values from a few firms (usually *not* representative firms) as proxy values for others may produce results that are not reliable. Furthermore, many other categories (even those with direct effects) will not have been studied. Additionally, some benefits, such as comfort, are difficult to conceptualize and make the category hard to measure. These cases, present the problem of generating consistent, reliable, defensible, and practical ways of valuing benefits that are brought about by program interventions.

We have examined a number of different approaches and evaluated them with respect to: 1) credible methods / demonstrated in literature; 2) ease of response by respondent / comprehension of the question by respondents; 3) reliability of the results / volatility; 4)

---

participant would realize 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 NEBs may be appropriately credited (as should the savings) during the interim. However, these are fine points around the principles discussed above.

[8] And if societal water benefits are relevant (e.g. a city is nearing infrastructure capacity) then marginal costs or costs for the "next" facility or other valuations depending on the local conditions may be applied for NEBs for that perspective.

conservative / consistent results; 5) computation clarity, and other criteria. We have tested and refined several approaches, such as:

- Willingness to pay (WTP) / willingness to accept (WTA) / contingent valuation (CV)
- Alternative methods of comparative or relative valuations
- Direct computations of value to owner
- Ordered logit, and other approaches.

These measurement methods can be complex to implement, and we have worked hard to refine the techniques to provide consistent and defensible results. These techniques have been applied to all of participant benefits categories that are listed in Table 1.

Values and Consistency from Alternative Measurement Methods

One of the difficulties of measuring participant side benefits is that data on participant impacts are not available in the literature, or even generally (or practically) from any method of direct physical measurement. Surveys generally provide the best source for the types of information that are needed to estimate participant impacts and benefits. As mentioned, because these benefits had hardly been measured before,[9] our research has explored multiple measurement methods to value these participant NEBs; ultimately, allowing us to see if different methods lead to consistent and reasonable estimates. The following sections focus on methods for measuring the hardest-to-measure types of benefits – participant benefits. Moreover they compare and contrast quantitative results from several distinct measurement methods, including:

- willingness to pay, willingness to accept, and contingent valuation, and
- relative valuation, and scaling techniques adapted from academic literature.

Note that our assessment focuses on comparisons of the value for the total set of NEBs discussed with the participant, not on the values associated with individual benefits categories. While we have applied these measurement methods successfully to individual categories in several studies, the examples provided here focus on measurement of total participant NEBs. The measurement practices also apply the valuation of individual categories of benefits; however, if the extra questions added to permit the comparisons addressed in this paper had also been added for individual NEBs categories, the surveys would have been impossibly long and the results easily may have been unreliable due to respondent fatigue.

Evaluating the Impacts of Different Types of Non-Energy Benefits Valuation Questions

The results from this analysis are indicative only – the experimental situation was not perfect, and the questions about measures, free ridership, and individual NEBs category valuations were asked at the end of an already relatively lengthy survey. In addition the small sample of 100 did not allow us to select many subgroups, so in some cases, the respondents may have been influenced by earlier questions. However, it was a useful opportunity to examine

[9] Space considerations preclude a literature review. Most of the references listed include extensive literature reviews; however virtually no studies have conducted work to actually estimate participant-side NEBs until Skumatz 1997.

variations in responses based on different question approaches within the same program group, and to illustrate strengths and weaknesses of different approaches to provide suggestions on "best practices."

Willingness to Pay, Willingness to Accept and Contingent Valuation Questions

Willingness to pay (WTP) surveys are often used to estimate hard-to-measure benefits, such as recreational benefits. Respondents are asked his / her WTP for a service or benefit. Specifically, in our surveys, we asked participants to indicate how much s/he might be willing to pay for the range of non-energy benefits and some individual benefit categories that they received from the program. The results provide one estimate of the value of benefits (whether positive or negative) to participants. Although WTP can be a useful technique, we found that the questions can be difficult for some respondents to answer. We have tested both open-ended questions and versions requiring customers to choose from a list of prompted value responses.[10] We phrased the contingent valuation / willingness to pay questions in four separate ways and they yielded different results. The questions were presented as: 1) Willingness to pay for the benefits received; 2) willingness to accept in exchange for the benefits provided; 3) Randomized binary contingent valuation question, in which random groups are given a low or high valuation number, with a maximum of one follow-up value, and 4) Non-random, incrementing binary WTP questions in a series

The results are presented in Table 2. Initially, we asked how much respondents would be willing to pay, independent of the program, for the total of the NEBs that they accrued and reported. The average of their responses was $670.[11] As a follow up question, we asked how much respondents would require in exchange for our taking away the NEBs that they reported. Here, the average answer was $674.

We then asked a randomized binary CV question. We informed our interviewees that, although the exact price of the NEBs associated with the program were unknown, the prices were suspected to lie within a certain range. Then, without telling them the range, we randomly asked them whether they would be willing to pay either $150 or $1,000 for the NEBs that they accrued.[12] If we asked them the higher value and they said no, we followed up with the lower value. If we asked them the lower value and they said yes, we followed up with the higher value. And if they said no to the lower or yes to the higher value, we stopped the question.

Table 2. Valuations Provided Using Willingness to Pay / Willingness to Accept Questions

Question Type	Value of Total NEBs, annual
Willingness to Pay for Overall NEBs	$670
Willingness to Accept in exchange for total NEBs provided	$674
Randomized binary CV	$91
Incrementing binary WTP series	$238
Compare to: average energy bill savings	$240/year

[10] An example question might be: when you think about the (positive and negative) benefits from reduced shut off and reconnect incidents you received from the program, what is the maximum you might be willing to pay for these benefits…". They were asked unprompted, but many respondents could not answer the question and they were then asked whether their WTP fell between certain ranges of dollar amounts.

[11] This was after dropping the highest value, which was 10 times greater than the next highest and appeared to us to be an outlier. (The minimum values were all self-reported zeros, and therefore credibly *not* outliers).

[12] These values were determined following a small pre-test of WTP and WTA questions.

We asked a smaller pool of respondents the randomized binary CV question. The average "yes" response rate for the smaller value, $150, was about 21%. The average yes response rate for the $1,000 question was about 7%, though far fewer respondents were asked the question because of its conditional format.

Although more sophisticated techniques should be used to extrapolate from the responses to this binary question, some simple manipulation shows the difference between the two types of questions. Assuming that the 21% yes answer rate to the $150 question applied to the entire sample of 100 interviewees, and similarly for the 7% rate for the $1000 question, then the average valid WTP for the NEBs reported in our interviews would be close to (14*150 + 7*1000) / 100, or $91 (the 14 derives from subtracting the 7 people that would pay $1000 from the number of people who would pay $150 to avoid double counting). This figure is substantially smaller than the calculated sample average of approximately $670 from the self-reported WTP values. However, it was calculated from a small subgroup of respondents, which decreases the accuracy of the estimate.

Finally, we asked a series of nonrandom, incremental binary WTA questions that went from $25 to either $150; $500; $1000 or $1500. The distribution of yes responses mimicked the CV question yes distribution for values shared between the two; for example, the same seven people said yes to the $1000 WTP value. Yet, 40% of these participants reported that they would be willing to pay $150, rather than just 21% of participants who participated in the randomized question. Here we begin to see the bias in WTA questions relative to their WTP counterparts. Because the WTA values were incremental from $25 to $150, some respondents who would not have been willing to pay $150 for their NEBs were willing to accept $150 as compensation for having them (hypothetically) removed. Those responses that were logged as yes to the $150 WTP question answered "no" to the willingness-to-accept of $25 question, which indicates that their true willingness to pay lies somewhere on the open interval between $25 and $150.

Part of this difference may be attributable to the question format. It has been suggested (Cooper, Hanemann and Signorello 2001) that an incrementally increasing question format might lead interviewees to believe that they are in a bargaining situation, which could bias their responses. The expected direction of this bias is negative; however, in our sample WTP values were actually higher for the incrementally formatted question. Still, using the same technique as above (extrapolating the smaller pool means to the entire sample of $100), we obtain a WTP of about $238. This value is larger than the value generated from the randomized question, but it is substantially smaller than the value of $670 obtained from the direct valuation question.

When in theory the randomized question should have produced a higher estimate than the incremental question response, the most likely explanation for this disparity of bias direction is that we chose values that were not optimal. For example, it appears that 20% of those who were asked the binary questions would have been willing to pay $500 but not $1,000. For these respondents, the $1,000 value in the randomized question was too high, which biased their true WTP values downward.

In the end, these WTP / WTA / CV results are in accord with economic theory and the literature on contingent valuation. It is roundly accepted that exact, reported WTP values are given to bias, which is very apparent in our direct valuation questions that yielded average values far higher than any other question format.[13] Consumers are not accustomed to naming their own

[13] This is consistent with previous research Skumatz conducted in the residential sector. In that research, we found the responses to direct WTP questions were much higher in value than those estimated using comparative or other approaches. That former research did not investigate results from these refinements to the WTP / WTA / CV approach that are provided in this paper. See Skumatz, 2002.

prices, particularly for goods that they have never had to purchase in a traditional market situation (such as those in the social benefits category of NEBs). Further, when consumers do occasionally enter name-your-price situations, they often do so in the context of an auction or another bidding scenario, which can affect how they make purchasing decisions (WTP rises and consumer surplus drops to zero).

For our survey, the most efficient question format was the incrementally increasing binary value format. However, regression techniques may be able to extract information not obviously available from the responses to the randomized binary CV question, which appears to have biased WTP estimates downward as a consequence of aggressive boundary setting.

"Relative" or Scaling Questions

The other key method we applied in measuring the value of overall NEBs was comparative or scaling questions. Although the literature is full of references to the use of WTP in the measurement of HTM impacts (such as recreation), field experience made it clear to the authors that WTP questions were very difficult for respondents to answer – perhaps because the concept was hard for many respondents to understand. We noticed the problem especially with participants in low-income weatherization programs, but found similar difficulties with commercial and industrial participants as well. To accommodate for these problems the authors determined that if respondents were not able to assign a direct value (WTP), respondents might be able to indicate if the NEBs are more or less valuable than a marketable product or service. This simple idea turned out to have very large benefits in measuring NEBs.[14]

We used two key approaches in this study as well as previous studies and compared the results in Table 3. The approaches included the following.

- **Comparative or relative valuations**. We asked respondents to state whether benefits had a higher value or a lower value to them than energy savings – and to define in percentages how much higher or lower. We used this relative approach in previous projects for both residential and commercial programs and have found that it is very easy for respondents to answer (Skumatz and Dickerson, 1998, Skumatz et. al., 2000).
- **Verbal scaled valuations – modifications to the Labeled Magnitude Scale**. Based on work in psychology, biology, and the taste field, we asked several questions that jointly allowed us to determine quantitative estimates of the benefits' value. In particular, we asked about the importance of the non-energy benefits using specifically worded scale responses. Given previous research results from studies that used these scaled responses, we were able to assign relative numeric valuations for the NEBs and translate them into dollar valuations. The literature and research from these fields indicates that relative ratio scales found that the terms "barely detectable," "weak," "moderate," "strong," "very strong," and "strongest imaginable" had predictable, scalable relations to each other.[15]

[14] It is important to select an appropriate leverage value. For much of our work, we ask about value relative to energy savings, since it allows us to divorce issues of time period (per year, per month), it is something all participants know about generally and can relate to the program, and dollars can be assigned based on individual savings or averages based on evaluations. For other programs, we have had to select other items to compare to.

[15] This approach derives from work trying to rank differences in taste and pain sensations. Literature in this field (Green et. al. 1993) indicates that even though some adjectives do not have absolute meaning (e.g. "large mice" and "small elephants"), the relative distances among the adjectives remain roughly the same. The labeled magnitude scaling (LMS) work shows that the intensive adjectives that are often evenly spaced in categorical research are not evenly spaced on ratio scales.

We adapted this research to measure participant-side NEBS using carefully worded adjective scaled questions for various commercial / industrial and residential programs.

Certainly, the percentage questions can be difficult for respondents to answer – and a bit more time consuming. However, these answers provide more detail than verbal scales with responses such as "more valuable," which can be simpler and faster to answer. Concerns also arise about how to translate verbal scale responses into percentages or dollar values. The results collected for both of these methods are presented below. In commercial / industrial growth programs research that the authors conducted for NYSERDA and for a series of residential programs (Skumatz 2002), it was found that multipliers using verbal terms and numeric / percentage terms seem to coincide within about 10% of one another. We closely analyzed multipliers from the literature and compared it to our percentage terms and found very close matches – in fact as the sample size increases the multipliers and percentages become closer yet. We have developed multipliers based on extensive secondary research and our own experience from past projects. Thus, if time is a constraint, verbal multipliers may be a very viable substitute for more complicated percentage questions, assuming well-researched multipliers are applied.

The dollar savings predicted from using these multipliers are also an important component of the NEB valuation work.[16] For those cases in which energy savings is the value being compared to, the average (or even better, individual[17]) energy savings can be computed, and valued using the marginal tariff for the relevant customer class. In this case, the annual energy savings were estimated to be approximately $240.[18]

In principle, it might be argued that participant valuations of savings is the preferred multiplier, since that is the number respondents presumably had in mind as they answered the questions. In other cases, these data are not available, and we use the program computed savings value. We usually find that these savings estimates are similar in orders of magnitude, but may vary, depending on the program. Table 3 presents the resulting translations of the NEB values.

Table 3. Valuations Provided Using Comparative / Scaling Questions

Question Type	**Value of Total NEBs multiplier**	**NEB Value using Program estimated savings ($240/yr)**
Percentage value relative to energy savings	115%	$276
Translation of verbal scaling into multiplier terms	104%	$250

Comparing Value Results from WTP/WTA/CV Approaches with Relative/Scaling Questions

The results of these comparisons (and other work, including Skumatz 2002) indicate that the values for direct WTP / WTA generally produce much higher NEB estimates than results derived from the comparison methods. Using enhanced WTP approaches can reduce these differences and the results from these approaches provide lower NEB estimates.

[16] In some of our work, it has been appropriate to stop at this point in the valuation. For work for NYSERDA, the ultimate purpose of the work was for benefit cost ratios and tests. Rather than translate to dollars, the NEBs were included as add-ons to the energy savings, and valued and treated in this way throughout the benefit/cost valuation process. This eliminated issues of valuing energy one way for NEBs and another way for other stages of the B/C work. This approach was developed with Rick Ridge, Ridge and Associates, who was responsible for the B/C work.

[17] An analysis of the differences between applying individual multipliers to individual savings values vs. Applying averages will be analyzed in another paper at a future date.

[18] Determined using the weighted average of deemed savings and a marginal kWh rate of $.12.

Table 4. Comparison of Dollar Values of NEBs from Alternate Estimation Methods

Question Type	Value of Total NEBs, annual
Willingness to Pay for Overall NEBs	$670
Willingness to Accept in exchange for total NEBs provided	$674
Randomized binary CV	$91
Incrementing binary WTP series	$238
Comparative value using percentages	$276
Verbal comparative results	$250

Table 5. Adjustments to NEBs for Free Ridership to Derive "Attributable" NEBs

Adjustment Rationale	Value
Estimate of free ridership for the program	25%
Adjustment factor for NEBs for this program	25% reduction in $240 annual savings leads to estimated attributable NEBs of $187 / year per average participant using the verbal comparative results.

Over more than ten years of research experience performing these analyses, we have found that comparative or relative valuations[19] perform substantially better and more consistently than direct WTP methods. Willingness to pay (WTP) can often provide very volatile numbers because respondents find it difficult to understand the concept of stating or defining a dollar amount that they would be willing to pay for these benefits. We have incorporated multiple measurement methods into the same studies, and have found that on average, WTP is more volatile (and less conservative) than other approaches.

Not only are the responses more conservative and less volatile – relative valuations provide a superior measurement method in terms of eliciting responses.[20] Respondents can readily answer whether these other benefits are more valuable or less valuable than energy savings or another benchmark.[21] This is important. If respondents have to consider their response for a long time the answers they give are likely to be guesses. Our field experience with thousands of interviews and surveys has shown that this is not a problem for the relative responses – most respondents can talk (at length) about these answers. In addition, asking value relative to savings helps eliminate the time aspect of the values – we do not have to worry about figuring out what discount rates they may be using if we ask total value or some other valuation approach. Using relative responses tends to reduce the number of "outliers" who generally have no idea where to start with WTP responses.

As mentioned before, non-residential respondents occasionally have information on the direct value of some of the benefits. Direct respondents' valuations of benefits (based on internal studies) are also useful methods of measuring impacts, but many respondents have not conducted these studies, or have conducted them on only a few of the benefit categories. This produces too many missing values to be useful, and the available values are usually not representative. Some researchers have also used detailed pre-post and statistical methods of measuring impact values. These are very useful and promising; however, the approach is usually applicable to only a few impact categories (e.g. productivity), which still leaves many important benefits unmeasured.

We are continuing to research refinements for the questions on current and upcoming commercial / industrial and residential programs, which includes more work on WTP approaches and applications of ranking and ordered logit approaches to these measurements.

---

[19] Methods pioneered and adapted by the authors, based on the academic literature; descriptions in Skumatz (2002).

[20] Which is also likely a reason the results are more reliable and less volatile.

[21] This was a key development – this lever or relative approach (see Skumatz 1997). Asking value relative to an item for which a dollar value can be assigned provides a method to compute the total value of the NEB. The lever we use varies by program.

NEBs Results in Context

Table 6 provides overall findings related to the magnitude of NEBs from residential and non-residential programs. NEBs are hard to measure, which has led to exclusion of these benefits from program evaluation and the decisions that derive from the evaluations. However, that practice leads to the implicit use of the value of zero for the NEB. The results below show how that simplification can be detrimental.

Note that one of the advantages of using comparative valuation methods over WTP is that the results are more easily compared across jurisdictions. The results can help normalize for areas with large heating / cooling loads compared to areas without these seasonal loads. We have compared results for various areas and programs and noticed some patterns. These values are presented in terms of relation to energy savings, because dollar values would vary based on the size of the program / measures (we believe this provides a more consistent basis for comparison).

Certainly the results and currency valuations vary by type of program, climate, measures included, and program targets. The results for the measure-based program used in these demonstration computations indicate attributable participant benefits equal to about 75% of energy savings, after adjusting for free ridership. We find that including gas measures can increase the NEBs associated with the program. Comfort-related NEBs are higher (especially in currency value) in climates with higher heating (or cooling) needs. However, the table provides some information on the range of omitted NEB values. Given these values, even including participant valuations of NEBs in a benefit - cost analysis, improves the benefit-cost performance by about one-third, depending on the program type.

Table 6. Orders of Magnitude Results/ Rules of Thumb for Residential and Non-Residential NEBs[22]

	NEB Values Relative to the Value of the Program Energy Savings (easily converted to currency values)
Perspective	**Non-residential Programs**
Utility	Perhaps 10% for non-low income programs; 30%-50% or more for low income programs; varies for commercial but tends to be relatively low unless power quality is an important factor. Varies with inclusion of gas measures or not, program design and targets.
Societal	100%-300% of energy savings typical, depending on aggressiveness of environmental and economic valuations
Non-residential Participant	Equipment programs, 35-50% of measure savings;[23] New construction programs about equal to value of energy savings; Commissioning, about equal to the value of the work conducted; Other variations by program type. Depending on program type, these may need to be discounted additionally for free ridership.
Total NEBs	Total - 35% to 300% / 400% of energy savings[24]

Uses of NEBs

Early NEBs research was applied to estimating savings for utilities, for example lower arrearages for residential programs. This expanded to the use of NEBs for improving benefit-cost analysis. However, NEBs are not only useful in assessing value from the program, but provide a

[22] This table is based on SERA research on more than 40 programs. See, for example, Skumatz (2001) and Skumatz (2002).
[23] Our research has developed detailed estimates depending on the measures and program design. In addition, the key / most valuable benefits have also been assessed, providing useful information for marketing and targeting.
[24] Depending on the aggressiveness of the valuation methods applied for economic and environmental benefits.

more sophisticated method of analyzing benefits and barriers,[25] and provide exceptional guidance for program targeting, marketing, and design. Applications include the following.

Benefit-cost analysis. Some share of these effects may be appropriate for inclusion in program benefit-cost analysis or regulatory tests -- generally subsets or pieces of the NEB analysis. For example, for a project in California the authors helped identify those categories of net NEBs that were most appropriate to be included in a revised public purpose test for low-income programs. The results of other analyses have been used to develop NEB adders in other states. Some clients have elected to incorporate a proportion of the estimated NEBs into benefit-cost analysis computations for evaluation and regulatory applications. While perhaps not the primary use of NEBs work, these measurement efforts assess (and value) another set of impacts of the programs beyond what would have been realized without the program. The individual categories of NEBs were not provided in this paper, but the results of the dollar NEBs valuations for each of these individual NEB categories varies based on the program's design, measures included, sectors, targets, regions, and other factors. Estimated properly, these represent attributable impacts.

However, as mentioned above, only a subset of these impacts may be appropriate for inclusion into a benefit-cost analysis or into specific regulatory tests – and that subset depends on the purpose of the test. For example, some of the societal benefits may belong in societal tests, and a number of the participant benefits may be appropriate for inclusion in tests for low income programs because these programs often have specific program goals of reductions in hardship and bill-payment improvements. Identifying the particular subset of NEBs to be included in a particular computation for a program depends on the application / use of the computation. NEBs that are important some computations are not appropriate for other computations. Rarely is it appropriate to include all the NEBs in a computation. On the other hand, perhaps it should be equally rare to include none of the NEBs in a computation of program effects.

Measuring barriers to adoption of programs. While most evaluations analyze barriers on a verbal or numeric scale (e.g. "very important", 1-5) using NEBs can help determine the value or cost associated with that barrier (a "negative" NEB in NEB terms). NEB estimates are far more specific and provide a more practical and useful metric for assessing the importance of barriers. This analysis can help focus attention on (1) more concrete information on the level of importance of particular barriers and (2) potentially, the level of rebate or program intervention or investment in program attention necessary to address particular barriers.

Program design and targeting. NEBs can be incorporated into initial decision-making. For instance, NEBs can be used when making a decision about measures / features to include in new / remodeled buildings. They can also be used when computing the costs and benefits of investments in energy using equipment. In addition, NEBs can be used to identify key measure or program barriers that need to be addressed. For example, if maintenance NEBs are negative for specific measures – which indicates a possible barrier – the program may benefit from: (1) changes in measures; (2) additional education; or (3) investigations to determine if greater infrastructure for spare parts is needed.[26] An example of how NEB estimates provide useful insight into barriers may be illustrated by an example in which noise is determined to be a significant negative NEB for lighting. In this case, further exploration of energy efficient

---

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

[26] The value of the negative NEB provides information on the approximate cost of this problem where it has arisen, and how much of a rebate or intervention might potentially be needed to offset this issue / barrier.

lighting equipment may be warranted; there may be a problem with the specific brand of equipment. The NEBs can also be used to select measures and target groups that will maximize the program impacts. This can maximize the program's effectiveness given a fixed budget.

Marketing / outreach for programs: The NEBs information can be used to refine program outreach. As mentioned above the information can help identify specific target groups that will benefit most from participation, and ultimately help guide outreach strategies. The work can also identify key messages or selling points; based on the NEB results, an effective product advertisement may not be based on energy efficiency. In fact, energy efficiency / conservation benefits are often not the highest valued benefits participants derive from programs. We find that commercial / industrial participants value other benefits such as fewer maintenance repairs, improved productivity, and fewer tenant complaints more. This may partly arise because businesses can be skeptical about actual energy savings since savings can washout for commercial buildings in the midst of all the other changes that occur from year to year. Hence advertisements that reflect the NEBs, which they receive regardless of the energy savings, may be more effective at increasing interest in program participation (Bicknell and Skumatz, 2004).

Conclusions

Even though non-energy benefits (NEBs) are often ignored, they are an important set of benefits that provide energy-related measures and features in residential and non–residential buildings. These NEBs can be measured, and they represent effects that are attributable to programs. Moreover, these NEBs can provide useful information to help guide program decision-making and planning. Overall the measurement approaches described in this paper provide consistent and defensible estimates of the relative NEBs recognized by participants.

Utilities may run energy conservation programs to reduce energy use, and builders may build homes and commercial buildings that include energy saving features, measures, and designs. However, energy savings may not be, and appear not to be, the highest valued outcome to buyers / participants. We have conducted research on NEBs for many commercial/industrial programs, and have analyzed the results by measure, building type, and program type. The results from this research have yielded useful feedback for program planners, evaluators, and regulators. Also, these NEB results have applications in benefit-cost analysis, regulatory tests, and program design. Yet, NEBs may be most valuable in their applications to marketing and outreach. Assessment of net non-energy benefits is critical to understanding the full range of benefits provided by programs.

References

Bartoshuk, Linda M, Valerie B. Duffy, Katharine Fast, Barry G. Green, and Derek J. Snyder. 2000. "Hormones, age, genes, and pathology: How do we assess variation in sensation and preference?" Danone Foundation Symposium Proceedings from "Food Selection, From Genes to Culture," Paris, France, December 2000.

Bicknell Charles, and Lisa A Skumatz. 2004. "Non-Energy Benefits in the Commercial Sector: Results from Hundreds of Buildings." Proceedings of the 2004 ACEEE Summer Study, California, USA, August 2004.

Cooper, Josesph C, Michael Hanemann, and Giovanni Signorello. 2001. "One-and-One Half Bound Dichotomous Choice Contingent Valuation." Department of Agricultural and Resource Economics, UCB, UCDARE, Working Paper 921. http://repositories.cdlib.org/are_ucb/921

Green, Barry G., G.S. Shaffer, & M.M. Gilmore. 1993. "A semantically-labeled magnitude of oral sensation with apparent ratio properties." *Chemical Senses*, 18, 683-702, 1993.

Hall, Nick, Lisa A Skumatz, and Lori Megdal. 2000. "Low Income Public Purpose Test: Non-Energy Benefits for Low Income Weatherization Programs." PG&E, 2000.

Imbierowicz, Karen and Lisa A Skumatz. 2004. "The Most Volatile Non-Energy Benefits (NEBs) - New Research Results 'Homing In' On Environmental And Economic Impacts." American Council for an Energy Efficient Economy Summer Study." ACEEE, Washington DC, USA, August 2004.

Pearson, Dennis and Lisa A Skumatz. 2002. "Non-Energy Benefits Including Productivity, Liability, Tenant Satisfaction, and Others – What Participant Surveys Tell Us About Designing and Marketing Commercial Programs." Proceedings of the 2002 ACEEE Conference, California, USA August 2002.

Skumatz, Lisa A. 1997. "Recognizing All Program Benefits: Estimating the Non-Energy Benefits of PG&E's Venture Partner's Pilot Program (VPP)." Chicago, Illinois: *Proceedings of the 1997 Energy Evaluation Conference,* Illinois, USA, 1997.

Skumatz, Lisa A. 2001. "Non-Energy Benefits (NEBs) – A Comprehensive Analysis And Modeling Of NEBs For Commercial & Residential Programs." Proceedings of the AESP Conference, Florida, USA, December 2001.

Skumatz, Lisa A. 2002. " Comparing Participant Valuation Results Using Three Advanced Survey Measurement Techniques – New Non-Energy Benefits (NEB) Computations of Participant Value." Proceedings of the 1998 ACEEE Conference, California, USA August 2002.

Skumatz, Lisa A, and Chris Ann Dickerson. 1998. "Extra! Extra! Non-Energy Benefits of Residential Programs Swamp Load Impacts!." Proceedings of the 1998 ACEEE Conference, California, USA, August 1998.

Skumatz, Lisa A, Chris Ann Dickerson, and Brian Coates. 2000. "Non-Energy Benefits in the Residential and Non-Residential Sectors – Innovative Measurements and Results for Participant Benefits." Proceedings of the 2000 ACEEE Summer Study, California, USA, August 2000.

Skumatz, Lisa A, John Gardner, and Charles B Bicknell. 2005. "Techniques for Getting the Most from an Evaluation: Review of Methods and Results for Attributing Progress, NEBs, Net to Gross, and Cost-Benefit." European Council for an Energy Efficient Economy Summer Study, Cote d'Azur, France, May 2005.