NEBs: The Latest in Results, Applications, and Best Practices for State Cost-Effectiveness Tests

Lisa A. Skumatz, Ph.D., Skumatz Economic Research Associates, Inc., Superior, CO

ABSTRACT

Utilities implement conservation (or energy efficiency) programs with the goal of reducing energy, or kWh, use. However, they have long ignored the array of indirect, or non-energy impacts, deriving from these programs. These benefits accrue to three classes of beneficiaries: participants (comfort, productivity, etc.); the utility (bill-payments, system reliability, etc.), and society (emissions, job creation, etc.). Non-energy benefits (NEBs) have been a topic of conversation for 20 years, and after a time being used mainly for program marketing and targeting, are being incorporated into benefit-cost tests. A variety of regulatory cost-effectiveness tests are used to compare the cost-effectiveness of programs and optimize program investment (TRC, UCT, RIM, others). This paper examines revisions to reduce bias and better guide investment between and among programs. Making improvements in the tests comes down to a question of the costs and benefits, and the associated accuracy, of the *improvements* in values or components (including NEBs). We examine the role of (subsets of) NEBs that are appropriate for inclusion to reduce possible bias in the cost-effectiveness tests, and current treatment of NEBs in cost-effectiveness tests and program / measure screening in states around the US. This paper provides information on best practices, and the progressing treatment of NEBs in benefit-cost tests in states across the country, including: limited "adders"; "easily measured NEBs"; hybrid and broader treatment of NEBs.

The paper reviews the range of values that 20 years of analysis are indicating as appropriate for a representative program type, and issues related to transferability of the results to other programs, and to benefit-cost work. Making improvements in the tests comes down to a question of the costs and benefits, and the associated accuracy, of the *improvements* in values or components (including NEBs). The paper assesses those elements of the benefit-cost equations with greatest uncertainty, and how that affects how evaluation expenditures should be optimized to best support decision-making. The paper examines the "bang for the buck" in uncertainty and risk associated with including NEBs metrics in decision-making. We discuss progress in NEBs measurement, and revisions to tests that are practical – and politic -- and decrease bias in screening tests.

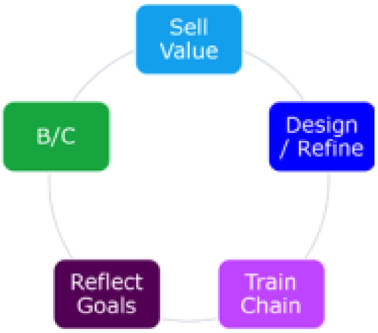
Introduction

For more than two decades now, researchers have recognized that energy efficiency programs deliver value beyond the intended energy savings, reflected in an array of other effects, called non-energy benefits (NEBs). These benefits accrue to the utility / agency offering the program (and its ratepayers) residential and commercial participants, and the broader region / society. Until recently, these (positive and negative) impacts have been largely relegated to use as program marketing tools, but more and more states are including measured NEB values or placeholder / proxy "adders" into program benefit-cost tests. This paper reviews the status of this evolution in the use of non-energy benefits in energy efficiency program planning, evaluation and decision-making.

Uses of NEBs

Non-energy benefits have five main uses, as illustrated in Figure 1.

Figure 1: Main Uses of NEBs



Source: Skumatz, 2014

Sell Value: They can be used in marketing, to help potential program participants recognize the value of the benefits they may receive from adopting the energy efficient measures, expanding their computations of return on investment (ROI) from participation from including only energy savings to include the broader list of net benefits associated with the measures.

Design / Refine: The NEB values can be used to help optimize program design and delivery. Given a fixed budget, the NEBs values can help identify the best mix of measures and/or program targeting to maximize program benefits or return. Groups with greatest return can be targeted, and the positive and negative NEBs can be extremely informative in setting the appropriate level of financial incentives to gain the participation desired, or to identify and offset any negative NEBs (aka "barriers") associated with the program or its measures.

Train the Chain: The program's delivery and success often depends on the performance of implementers, and an array of upstream actors for market transformation programs. Making sure these actors are cognizant of NEBs valued by potential participants – and delivered by the equipment included in the program – improves participation and measure uptake.

Reflect Goals: Low income programs have goals beyond energy savings (e.g. hardship goals, etc.), and some other programs do as well (e.g. job creation goals for ARRA programs). NEBs specifically reflect the value and performance toward an array of "non-energy" goals.

Benefit-Cost Test: Utilities and regulators use benefit-cost tests to make decisions about energy efficiency (EE) measures, programs, and portfolios. Including a fuller array of the net benefits from programs (estimated as NEBs) – beyond energy savings – reduces bias in decision-making and resource allocation regarding programs.

The remainder of the paper focuses on the uses related to benefit-cost tests (and, as an aside, the reflection of program goals).

The Bias in Benefit Cost Testing

In economics, benefit-cost tests used to inform decision-making would include the present value of the stream of attributable benefits, divided by the present value of the stream of attributable costs. The *costs* of energy efficiency programs are factored into the assessment of benefits and costs that are required steps in determining which energy efficiency programs are funded, but the list of benefits traditionally included in protocols is often limited to energy savings. NEBs may have been omitted because they were considered "hard to measure" at the time the protocols were developed, and no explicit placeholders appear to have been included to accommodate measurement progress. Leaving explicit place for all attributable costs and all attributable benefits would have allowed for improved evaluation methods and results – and allowed for less bias over time. The theoretical consistency of the test(s) can be best improved by:

- Including monetized estimates of the NEBs (net positive and negative) in the Total Resource Cost (TRC) test computation; or
- Excluding all NEBs and the costs associated with achieving the NEBs, or use the Utility Cost Test (UCT) that involves only the costs paid by the utility.

There are five major benefit-cost tests, and the incorporation of NEBs in the formulations are included in Figure 2 below.

Figure 2: Summary of Benefit-Cost Tests and Potential NEB-Based Updates (Skumatz *et. al.* 2009, Amann 2006, updated)

Test	**Benefits**	**Costs**	**States Using Traditionally**	**Improved treat-ment with NEBs**
Utility Cost (or Program Administrator Test) (UCT or PAC)	• Avoided supply costs for transmission, distribution, and generation (TD&G) • Avoided gas and water supply costs	• Program administration • Participant incentives • Increased supply cost	CA, CT, HI, IA, IL, IN, MI, MN, MO, NY, OR, RI, TX, VA, WA, BPA	Use cost only paid by the utility
Ratepayer Impact Measure (RIM) (or No Loser's Test, or non-participants test)	Same as above plus • increased revenue	Same as above plus • Decreased revenue	AR, CO, FL, GA, HI, IA, IN, MI, MN, NC, ND, NV, SC, VA, WI	
Participant cost	• Utility bill reductions • Participant incentives	• Participant direct costs	AR, CA, FL, HI, IA, IN, MI, MN, NY, VA	Participant NEBs
Total Resource Cost (TRC)	• Avoided supply costs for TD&G • Avoided gas and water supply costs • Utility bill reductions	• Program administration • Participant incentives • Participant direct costs • Increased supply costs • Decreased revenue	AR, CA, CT, CO, GA, HI, IA, ID, IN, MA, ME, MI, MO, MT, NH, NJ, NV, NY, RI, SC, UT, VA, WA	Include all participant and utility NEBs; (costs are already included)[1]
Societal / Societal Cost Test (SCT)	Same as above plus • Externality benefits (reduced pollution, improved reliability, etc.)	Same as above	AZ, IA, ME, MN, MO, MT, NJ, OR, VT, WI	Include all NEBs – utility, societal, and participant NEBs valued (costs included)

Chicken and Egg: Making Progress in NEB Values

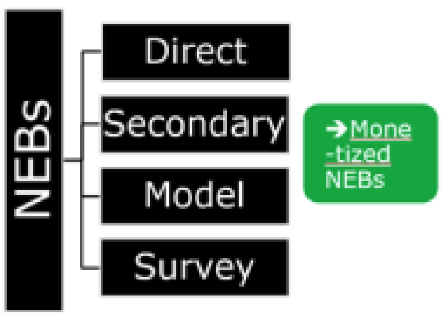
Although it was well-recognized that NEBs had an appropriate place in benefit-cost equations early on, the omission of NEBs from the equation may be understandable when values for the entries were unavailable. Limited research began in 1994 to estimate values for a variety of NEBs; however, funds for estimation of a large number of NEBs were not readily available, because NEBs did not have a clear place in important program decisions, like benefit-cost work (which might justify expenditures of funds). This twenty years of chicken and egg has only recently begun resolving. The twenty years from 1994 to the present allowed slow accumulation of a body of literature of NEBs estimation work, early on

[1] Another treatment would be to exclude all NEBs and all costs associated with achieving them (or use the UCT).

to be used in program marketing, and now becoming sufficient to support application to benefit cost work. A history of this progression has been summarized elsewhere (Skumatz 2014, Skumatz et al. 2009) reflecting progress from hypothesized lists of generalized benefits that might be attributable to programs, to tentative applications in low-income programs, to full-fledged estimation work of scores of categories of NEBs for literally hundreds of programs across the nation, and recent re-consideration of NEBs in cost-benefit applications.

The estimation of NEBs uses one of several approaches, depending on the type of NEB category being estimated. The point is to develop monetized estimates of NEBs; hand-waving, or "large / small" categorizations do not solve the underlying problem of getting NEBs taken seriously, and using them to reduce the bias in calculated benefit cost ratios. To date, best practice (and the bulk of the recent literature) for measuring individual NEB categories can be categorized into four main approaches, illustrated in Figure 3.

Figure 3: Monetizing NEBs



Source: SERA

Direct measurement: In some cases, NEBs can be directly measured, using records, statistical analysis on impacts, or other direct methods (e.g. arrearages,[2] sales impacts). Regressions or direct measurement are available for some applications; they have been used for a limited number of studies of productivity changes, and for disaggregating program level NEBs to individual measures (reviewed in Skumatz 2009).

- **Secondary data**: A substantial number of NEBs categories can be estimated or bounded using information from secondary literature. The calculation is based on the program-attributable *change* times the *value* of that change, with one or more of these inputs derived from secondary data (Examples include: e.g. reduced fire incidence from safety improvements times insurance values from tables providing the damage to life and property from fires,[3] water savings estimated using average usage from water association data and customer rates from national tables, value of changes from fewer bill-related phone calls, etc.).[4]
- **Models**: Affordable, peer-reviewed third party models are available and have been particularly applied to estimating NEBs for two major categories: emissions and job creation / economic impacts. The key analytical issue is to measure net impacts, estimating impacts compared to an appropriate base case.[5] Emissions valuations can be derived from markets or regulatory proceedings or other appropriate sources.

[2] Many arrearage studies were commissioned by utilities, especially in the 1990s. Specialized statistical arrearage studies are used to estimate carrying costs for bad debt, reduced phone calls, reduced shut-offs, and similar impacts (this has been most commonly applied to low income programs).
[3] These approaches were demonstrated in several early studies cited in Skumatz 2002, including Brown (1994), Magouirk (1995), Skumatz (1996) and others.
[4] Appropriate valuations for saved utility staff time might be the marginal wage rate; household time would be valued differently. Similar approaches can be applied / adapted for O&M, and some other categories.
[5] The important best practice is to measure "net" impacts of the new "savings scenario" compared to a base case. For example, best practice is either 1) assuming the jobs associated with electricity generation vs. jobs associated with the savings case, or 2) assuming jobs from the program (and savings) case, vs. a base of jobs associated with the program funds instead being spent on a consumer market basket of goods (*a la* consumer price index), arguing that program funds are raised from a

Survey-based Monetization: Estimation of many important participant-side NEBs like comfort are best estimated using specialized surveys. The methods that are now most common industry standard were pioneered in the late 1990s and early 2000s (Skumatz 2002 for summary), and are based on academic methods that ask participants to rank or value the NEBs relative to other services derived from the program and measures, rather than relying on willingness to pay or other approaches, to develop monetized NEBs.[6]

Deeper reviews of the measurement approaches are provided in other reports (Skumatz 2010b, Skumatz 2009, TekmarketWorks and Skumatz and Megdal 2001). The key, however, is to use approaches that allow monetization of the NEBs so they may be included directly in benefit-cost applications and the other uses mentioned in this paper. To do less ignores the progress of 20 years and reduces the uses to which the values can be applied.

Value Ranges for NEBs

Overall, based on results from NEBs literature over the years, benefits to the utility / agency are the lowest in total value, and those occurring to society and the participants are quite a bit higher. This is summarized in Figure 4, using the case of low income weatherization programs as illustration.

Figure 4: Ranges and Typical Values for NEBs for One Program Type

Range for Total Utility, Societal, and Participant NEBs

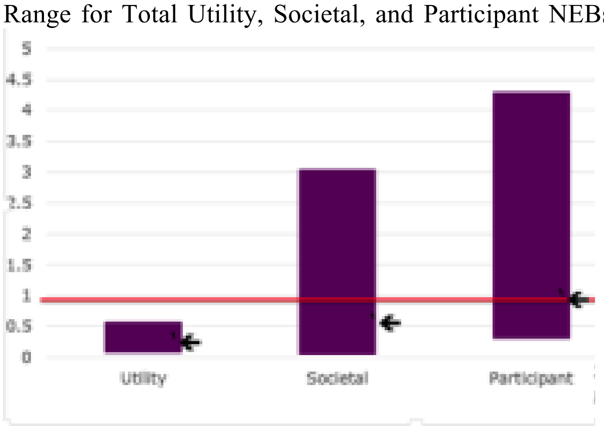


Figure 4 presents a typical range for NEB values for Low Income Weatherization Programs, expressed as a multiplier relative to the energy savings associated with the program. The largest constituent NEBs in each group (beneficiary category) are listed below.[7]

- Utility NEBs: low income subsidy avoided; arrearages / payments.
- Societal NEBs: Economic / job impacts; water savings; emissions.
- Participant NEBs: comfort, noise, O&M, water bill savings, home improvements; health and safety. Important participant NEBs for commercial programs include tenant satisfaction, productivity, O&M, others.

There are scores of NEBs studies of low income, residential, and commercial programs conducted for utilities and agencies around the country. Examination of the values for NEBs from the body of literature has identified patterns and similarities in values for many NEB categories; however, even when normalized to units of "per participant per year", the participant and utility / agency NEB values show wide ranges, because the specific values depend on:

- the sector to whom the program is offered (residential, commercial, etc.)
- the measures included in the program (one appliance vs. whole home weatherization, etc.),

system benefit charge charged to customers who would otherwise have spent the funds on their standard bundle of goods. (see Skumatz et al 2009).

[6] These survey approaches rely mostly on labeled magnitude scaling or "relative" approaches rather than willingness to pay or other direct approaches, for factors like comfort, noise reduction, quality of light, and other factors. These can also be used to provide estimates for a number of additional commercial benefits, if other approaches are not available (e.g. tenant complaints, possibly productivity, etc.).

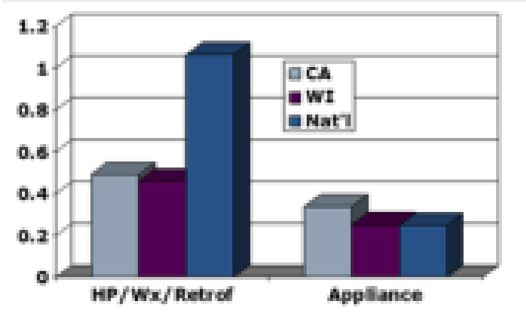
[7] Results derived from data on NEBs from detailed data from national studies included in "NEB-It" model (SERA, 2000-2015)

- the climate zone (for programs with HVAC measures),

Other determinants also matter in a somewhat more secondary manner: the type of program (rebate vs. financing vs. design assistance, etc.), the demographics or business types included (targeting homes with elderly or sick have higher values for comfort; low income programs may have larger payment improvement benefit to utilities, etc.), and other factors.

Some NEBs are fairly program-independent. The most straightforward example of this is emissions. Avoided emissions are directly related to the generation mix for the utility. While emissions may vary based on the degree to which the program savings affect base load vs. peak generation, there is a fairly direct relationship.[8] In a fairly simple way, factors – in dollar or percentage terms – can be applied to account for the emission effects, and these can be attributed directly as an "adder" benefit in direct ratio to energy savings.

Figure 5: Relative Economic / Job Creation NEBs for Two Program Types
Source: Skumatz and Gardner 2007



Other NEBs do not have such a direct relationship. Economic / job creation effects vary dramatically based on the type of program, the territory under consideration, and local industry mix in the territory. An example of results estimated by the author follows in Figure 5. The figure shows that relative multipliers for net job creation and economic impacts are larger for weatherization programs than for appliance replacement programs (higher bars for each of three territories). In addition, the low bar for national effects for the appliance program vs. the high one for the retrofit program indicates that insulation or some of the components installed in weatherization are made somewhere in the US; appliances do not seem to be. Thus, a million dollars spent for one program vs. another have big differences in terms of their economic impacts at the local and national level.

Best Practices for NEB Values: It is possible to adapt some NEB values from other jurisdictions, but there are key, high-value NEBs that vary based on program type, measures included, and sector (residential vs. commercial). Examples of those that may be adapted in fairly straightforward ways include ratios for T&D losses, water savings, and some others. However, emissions depend on local generation mix (and potential air sheds, etc.), economic impacts vary with program, measure, and local job mix; and many others also vary.

Which States are Where?

States are increasingly considering incorporating NEBs into benefit-cost tests, but to deal with the cost-and-accuracy issue, states are taking one of several approaches:

- Incorporating a simple, conservative "adder" to the benefits. Most suggest they are trying to incorporate factors related to omitted environmental or emissions effects (emissions effects are probably especially suited to incorporation as adders). In some cases, these "adders" represent conservative placeholders for the broader concept of NEBs. Some states add an extra premium or adder for low income programs.

[8] Certainly, this is a simplification. The quality and vulnerability of the air shed and time of day / time of year affects the resulting health effects and ozone creation, etc.

- Incorporating "easy to measure" NEBs into the benefits. Several states are adopting this flexible approach, with the "easy to measure" benefits varying among programs (for example, including easier-to-measure water bill savings from clothes washer programs, and omitting "softer" NEBs such as comfort, measured from surveys);
- Trying to measure all NEBs, or the leading from among several dozen NEBs; or
- A hybrid approach: using an adder and measuring either easy-to-measure benefits, or as many benefits as possible beyond what is included in the adder.

Our recent comparison of states in terms of the extent to which they are considering NEBs is presented in Figure 6. The status of NEBs in these jurisdictions is constantly changing.

Figure 6: Examples of NEB Treatment in Regulatory Environment

Application	Utilities / regions	MORE AGGRESSIVE ==>
Program marketing	Fairly widespread use in utilities / states across the country	
Test / program screen: Adder	Currently, multiplier values in various states range from a low of 7.5% for gas and 10% for electricity, to 15% multipliers, sometimes adding another 15-25% for low income programs. Dollar adders range are currently $15 (carbon adders).	
Test / program screen: Readily measured	Examples of NEBs currently classified as "readily measured" or "reliable and with real economic value" in various states include: utility NEBs, property improvements, health and safety, comfort; costs of complying with foreseeable environmental regulations, maintenance, equipment replacement, greenhouse gas reductions or carbon values, measure lifetime effects, product loss, productivity, value of changes in floorspace from new measures, and some specialized low income program effects.	
Test / hybrid (potential adder and measured)	Currently, several states include an adder (above) and allow additional valuation of subsets of NEBs, including, for example, those with current market values, or water / sewer savings, present value of deferred plant extension, or other NEBs.	
Test / program screen: Broad	One state is incorporating broader NEBs if measured with surveys. Broad inclusion of all NEBs in official screening is not yet found.	

Source: For detailed state-by-state values see Skumatz et. al. 2009, updated Skumatz 2014

Traditional tests can be modified to include NEBs (as shown above), and benefit-cost tests can be crafted to especially recognize the value contributed by NEBs. The first benefit-cost test specifically developed to incorporate non-energy benefits as a policy element was the Low Income Public Purpose Test (LIPPT), developed for the State of California's low income programs (TekMarketWorks and Skumatz Economic Research Associates, 2001). This test incorporated a set of participant, utility, and societal values to represent the broader policy benefits of the state's low income programs, which had goals related to energy savings and additional policy goals. The values from this test have been in use in association with low income programs in California since that time. A number of years later, additional specialized tests were proposed (Amann 2006), and, another effort at refined tests (called the "Resource Value Framework") has recently been undertaken by NEEP (Woolfe, et.al. 2014). Whether the basic tests are refined, or specialized tests are developed for some subsets of programs, the basic principle remains the same as stated in the 2000 work; identify the goals, and refine the metric to best attribute and represent progress and support decision-making regarding the program(s) with a minimum of bias.

Best Practice for B/C Treatment: Best practice for values incorporated into benefit-cost tests is the case that reduces bias and includes the most comprehensive stream of (appropriate) NEBs in addition to the energy savings as a program benefit (as specified in Figure 2). Perhaps the best balance of practicality and bias reduction would be represented by a NEB value that uses a hybrid approach (approach recommended in Skumatz 2014, 2015a). This would be calculated as the combination of two factors:

- A dollar or percentage "adder" that represents those NEB values that are invariant with respect to program type (largely emissions, which are directly related to energy generation displaced, and might vary only by whether a program is largely "baseload" or "peak"). Other NEBs that depend mostly on the energy saved and not the measures or sector would also be appropriate to include (e.g. possibly T&D savings, etc.).
- Dollar or percentage values based on estimation of other NEBs with the highest values,[9] selected based on the sector and measures included in the program.[10] Some of these may be adapted from other studies (with similar measures, customers, climate, etc.). The "adder" approach – or adoption of values from other locales - may be able to be extended to climate independent measures and NEBs (for example, maybe lighting, which is relatively, but not completely, independent of climate). Those with a greater climate variation or business category variation may be best addressed by local measurement.

Given the current status of treatment by states around the country, best existing practice, in decreasing order, is represented by:

- Those with the most comprehensive set of NEBs allowed - Massachusetts,
- This is followed by the states that allow measured NEBs or an adder plus additional measured NEBs (BC Hydro, Oregon, DC,[11] Rhode Island, and, possibly Colorado and Vermont). Colorado and Vermont also incorporate relatively high adders.
- Several others allows measured NEBs for Low income programs, and may be best practice, depending on which NEBs are estimated and included

Higher adders (net of negative NEBs) or inclusion of more NEBs categories also represent progress toward bias reduction.

Transferability and Dominos

As mentioned, the literature includes scores of NEB studies on a host of residential, commercial, low income, new construction, and other programs. Patterns exist, and states (and utilities) would love to be able to take values from one study (or from sets of existing studies) and apply them more widely. The transferability is related to the factors noted above (climate, measures, etc.), but is also limited by the fact that the NEB categories estimated in previous studies varied because the clients for whom the studies were conducted often limited the number of NEB categories that they wished to have estimated; the studies don't have consistent NEB lists, so they need to be examined carefully in order to detect patterns and transferability.

However, states are leaning on each other. A recent study (Malmgren and Skumatz 2014) explored the domino effect for state incorporation of NEBs into benefit-cost tests. This study noted that NYSERDA and Colorado's early statistical work influenced adoption of adders for Vermont and DC and helped introduce more explicit and aggressive adders for standard and low income programs. As long ago as 2001, California constructed a low income public purpose test, incorporating NEBs to reflect the fact that program goals in the sector go beyond simple energy savings (Skumatz 2001). More recently, in the mid-2000s, NYSERDA developed estimates of a wide range of NEBs, considering the values and presenting scenarios including and excluding them, but they were not part of the official screening. This showed that NEBs could have a role in influencing decision-making. In 2008 and 2011,

[9] Possibly including others that are very easily measured (household water savings for example, depending on the program)
[10] NEBs with highest value can easily be identified from the array of existing studies (see Skumatz 2009, 2014 and others).
[11] DC is one of the few that explicitly requests values for green collar jobs and that would represent a best practice; discussed later in this paper.

interveners in Colorado led the State to undertake research (Skumatz 2010a) [12] and adopt an NEB adder for electric and low-income electric programs for its modified TRC test; Colorado revisited its cost-effectiveness screening in a later proceeding which resulted in an increase and expansion of the NEB adder to 25% for low-income programs for both gas and electric energy ratepayers.

Three elements contributed to the incorporation of cost-effectiveness screening in Vermont: the quantity and nature of available research regarding NEBs, a growing number of other jurisdictions incorporating NEBs into their cost-effectiveness screening, and the collaboration of a broad group of stakeholders. Vermont had a 4% adder to account for environmental externalities, and a 10% adder to account for reduced risk associated with obtaining electrical efficiency savings as an alternative to generation. In 2009, the Vermont Public Service Board revisited these cost-effectiveness screening values and cited the Colorado's review of its NEB adder in 2011. The Board ordered a 15% NEB adder as a conservative rebuttal presumption that could be amended in later proceedings, as information became available to more accurately assess the exact value of NEBs in energy efficiency in Vermont.

In 2009, the District of Columbia undertook consideration of NEBs, deciding that that the cost-effectiveness screening should involve NEBs "…including comfort, noise reduction, aesthetics, health and safety, ease of selling / leasing home or building, improved occupant productivity, reduced work absences due to reduced illnesses, ability to stay in home / avoided moves, and macroeconomic benefits," (DDOE-VEIC Contract, 2010). These benefits are included as a 10% NEB adder in the event that calculating NEBs requires significant original research. In addition to NEBs, the DCSEU's cost-effectiveness screening also involves a 10% risk adder, and a 10% adder for the reduction of environmental externalities. Although the inclusion of these adjustments for cost-effectiveness screening are important, the DCSEU stands out because of its incorporation of what has traditionally been seen as an NEB into performance benchmarks in six areas (energy consumption, renewable goals, reduced peak, low income targets, demand reduction, and development of green collar jobs).

Discount Rates

Questions have arisen around the specific calculation of benefit-cost ratios, even when the stream of costs and benefits are more fully identified. One is the choice of appropriate discount rates. We interviewed a number of states to uncover the sources and values of discount rates used.

The economic underpinnings of discount rates suggest that the appropriate discount rate is one that reflects the relative risk associated with the investment (Skumatz 2015a, Skumatz 2015b for this section). Our interviews found that there are three main discount rates that have been applied by include:

- the Utility's Weighted Average Cost of Capital (WACC). This depends on the individual utility's cost of debt (which is not treasury-based), cost of equity, and the capital structure. Values tend to range between 5 and 8% (or more) [13].
- a societal discount rate. The societal value used in Vermont is 3%, for example.
- the 10 year treasury or prime rate. 10 year Treasury returns have had a range from about 0.5% to 2%.

The WACC reflects the risk associated with the utility's investments, mostly in generation, transmission, and distribution assets, The relative risk for investment in energy efficiency programs (at least mature "widget" programs), would potentially be expected to be less than a utility's investment in

[12] The study's results (drawing on approaches from the 1990s and early 2000s in California, from the work the consultant had done estimating the NEBs for NYSERDA, and by other work) developed estimates for Xcel's low income programs (Skumatz 2010a). The adopted 25% value was an increase over previous values, but considerably below the estimated values. In 2013 testimony, the utility argued that the study's estimated values were "too liberal."

[13] From author-conducted interviews. Low end of range for return on treasury from Woolf, "Survey of Energy Efficiency Screening Practices in the Northeast and Mid-Atlantic", for NEEP, 2013. Author interviews indicate the low end from this study (5%) is not common and that the high end of the WACC range presented and beyond is more common.

generation plants, which can have substantial risks of delays, cost overruns in labor, materials, or unknown technology. Energy efficiency programs generally can be funded from expenses, and tend to occur in very controllable ways. This would argue that efficiency programs should have lower discount rates than generation resources (which usually use WACC). Social discount rates have been argued; the lower threshold for this investment might be the treasury returns, which could be considered an appropriate return for an investment of little to no risk. Analyzing in regulatory environments where utilities are rate-basing and not rate-basing the costs, the conclusions remains the same. In the prior case, the risk of recovery is lower for energy efficiency programs than generation, and in the later, the risk from investments from working capital do not need to reflect the capital-based risk embedded in the WACC.

Best Practice for Discount Rates: The discount rate used should reflect the risk associated with the investment. WACC reflects investment in power plants and utility; the appropriate discount rate for energy efficiency programs or low income programs in particular should be lower than the WACC (reflecting less risk) and closer to the social discount rate or toward treasury returns, depending on the conditions mentioned above.

How Accurate is Good Enough for NEBs? How Much to Spend?

Improving the tests comes down to the costs and benefits, and the associated accuracy, of the *improvements* in values or components. Utilities and regulatory agencies struggle with how to achieve that balance. What additional NEB categories can be accurately estimated within a reasonable evaluation budget? Practically, the question should be three-fold: (1) Which NEB categories are most valuable? (2) What value range arises from its reasonable-cost measurement? and (3) Does the inclusion of the high vs. low end of this *range* of values change the benefit-cost conclusion (to include or exclude the program or measure)? If the inclusion of the high and low end of the ranges results in different benefit-cost decisions, more money might be invested in refining the measurement calculation (assuming the program investment decision is valuable, and up to just shy of the value of that potentially wrong decision) *(*Skumatz et.al. 2009*)*. Investment should focus on those NEBs with greatest low / high variation. If the use of the high and low value of the NEB range results in the same program decision, no further investment to improve the estimate is justified. Investing a lot of money to refine estimates of a low-value NEB is money less well spent than refining estimates for a high-value NEB. These concepts are a type of value-based decision-making that is basic to most economists.

Another benchmark for how accurate NEBs may need to be is to look at the accuracy levels for other elements of the benefit-cost calculation. B/C tests, at a high level, are calculated as:

- A present value of a stream of benefits (traditionally annual energy savings), with the stream extending over a number of years based on measure lifetimes, discounted by an appropriate discount rate;
- Divided by the incremental costs associated with the measures (specific cost elements were noted in Figure 1).

Energy savings are often and repeatedly measured using impact evaluations. They are presumably fairly accurate. The next element is measure lifetimes. A review of measure lifetimes (Skumatz et. al. 2009) finds that 1) lifetime values for individual measures can vary widely between states (a factor of two for some measures), and 2) the values that are used are rarely based on statistical underpinnings. In addition, note that the values are based on the values have been in place for more than 20 years in many cases, and even if they were accurate, they may no longer be appropriate for the newer mechanics and decay behavior associated with new technologies being installed. The next element is discount rates. Discount rate values can also vary between states – from about 1% to 8% depending on the protocols (a factor of 8 or more). The last factor is incremental cost data. Accurate and appropriate incremental cost data are notoriously hard to come by and expensive to obtain. This comparison is provided to put the accuracy needs for NEBs into perspective; other key elements of the B/C ratio are not "written in stone" either, but have not received with as much skepticism or concern as NEBs.

Best Practices in Effort for Estimating NEBs: Measure the highest value NEBs, many of which can be estimated at low cost. Consider minimizing the incremental cost for NEB studies by routinely including a NEBs survey battery with process or impact evaluation surveys for programs. Use simplified multipliers or values from other studies (for similar programs), and develop ranges for high value NEBs (easily identified from the extensive NEBs literature). Develop a high and low range, and if the B/C ratio is above (or below) 1 when both the high and low values for NEBs are included in the calculation, the NEBs may be well-enough estimated. If not, additional research should be conducted, but not to exceed the "value" of a wrong decision about the program's omission or inclusion in the utility's portfolio.

Lessons and Conclusions

With more than 20 years of literature, NEBs have been subjected to deep research, the development and application of methodologies, testing in different programs, sectors, and regions, presentation at formal conferences, and associated peer and academic review. In particular, this evolution and development of the literature has moved the perception of NEBs among program staff, administrators, and regulators from that of general unfamiliarity and skepticism; to acknowledgement that some NEBs are "real," measureable, and useful. Stakeholders now have a larger appetite for considering NEBs in uses beyond "soft" applications like marketing, and for re-considering their use in benefit-cost applications. Twenty years of research has helped bolster the case that NEBs may be appropriate for additional applications; their values, although perhaps "hard to measure", are coming to be recognized as non-zero, and that begins to suggest that the regulatory tests in use across the country may be leading to biased programmatic decision-making.

Twenty years on, it appears to be time to reconsider benefit-cost tests that better represent actual benefits and costs, and support more optimal program investment. It is clear that there has been incremental progress in addressing the bias inherent in tests that exclude NEBs. In fact, we see a domino effect: as one state makes progress, another directly incorporates that progress into its deliberations. These case studies illustrate how New York's research informed Colorado's decision, which in turn influenced Vermont's decision.

The NEBs perspectives (societal, participant, and utility) overlap all other tests, and if a true representation of costs and benefits is desired, elements of NEBs represent exactly the missing factors in the benefit-cost calculation that reduce bias and better guide investment between and among programs and portfolios. NEBs analysis reduces bias in benefit-cost tests. The practical discussion usually centers on enhancing versions of the TRC (societal) test, given its broad scope. The TRC generally compares benefits in terms of avoided energy costs against program costs (involving both utility and participant costs). The paper provided best practices for NEBs estimation method, state treatments, discount rates, and other NEBs topics.

Billions of dollars are invested nationwide each year in energy efficiency programs. To accurately account for both the costs and the benefits of these programs, NEBs must be counted in cost-effectiveness tests. Society, ratepayers, and utilities will benefit by including NEBs in program evaluations, thus reducing bias in determining program cost-effectiveness. Inclusion of some NEBs is better than exclusion of NEBs, but long-term progress in addressing the bias in tests should not be delayed. Value-based decision-making can address the short-term measurement questions. We look forward to the next twenty years!

REFERENCES

Amann, Jennifer, 2006. "Valuation of Non-Energy Benefits to Determine Cost-Effectiveness of whole-House Retrofit Programs: A Literature Review", ACEEE Report Number A061.

Daykin, E. J. Aiona, and B. Hedman. 2010. *Picking a Standard: Implications of Differing TRC Requirements.* The Cadmus Group. http://www.cadmusgroup.com/wp-content/uploads/2012/11/Picking-a-Standard.pdf

Malmgren, Ingrid, and Lisa A. Skumatz, Ph.D., 2014, "Lessons from the Field: Practical Applications for Incorporating Non-Energy Benefits into Cost-Effectiveness Screening", Proceedings of the ACEEE Summer Study on Buildings, Asilomar, CA, August.

Skumatz, Lisa A., Ph.D., 2015a. "Testimony of Lisa Skumatz on Behalf of the Natural Resources Defense Council (NRDC), National Consumer Law Center (NCLC) and The California Housing Partnership (CHPC)", before the California PUC, Rulemaking A.14-11-007 et.al. April.

Skumatz, Lisa A., Ph.D., 2015b. "Non-Energy Benefits: Remaining Questions and Best Practices", Skumatz Economic Research Associates, Superior, CO, March.

Skumatz, Lisa A., Ph.D., 2014. "Non-Energy Benefits / Non-Energy Impacts (NEBs/NEIs) and Their Role & Values in Cost-Effectiveness Tests: State of Maryland", prepared for NRDC, Skumatz Economic Research Associates, Inc. (SERA), March.

Skumatz, Lisa A., Ph.D., 2010a. *Non-Energy Benefits Analysis for Xcel Energy's Low Income Energy Efficiency Programs.* Skumatz Economic Research Associates, Inc. (SERA), prepared for Xcel Energy, May.

Skumatz, L., 2010b. *Non-Energy Benefits: Status, Findings, Next Steps, and Implications for Low-Income Program Analyses in California.* Skumatz Economic Research Associates, Inc. (SERA), prepared for Sempra Utilities, May.
http://www.liob.org/docs/LIEE%20Non-Energy%20Benefits%20Revised%20Report.pdf

Skumatz, L. 2002. "Comparing Participant Valuation Results using Three Advanced Survey Measurement Techniques: New Non-Energy Benefits Computations of Participant Value," *Proceedings of the 2002 ACEEE Summer Study on Buildings*. Asilomar, CA. August.

Skumatz, Lisa A., Ph.D. and C.A. Dickerson. 1998. "Extra! Extra! Non-Energy Benefits of Residential Programs Swamp Load Impacts!" *Proceedings of the 1998 ACEEE Summer Study on Buildings*, Asilomar, CA, August.

Skumatz, Lisa A., Ph.D., and John Gardner, 2007. "Economic Impacts From Energy Efficiency Programs –Variations In Multiplier Effects By Program Type And Region", *Proceedings for the European Council for an Energy Efficient Economy (ECEEE)*, France, June.

Skumatz, Lisa A., Ph.D., Khawaja, M.S., and Colby, J. 2009. "Lessons Learned and Next Steps in Energy Efficiency Measurement and Attribution: Energy Savings, Net to Gross, Non-Energy

Benefits, and Persistence of Energy Efficiency Behavior." For CIEE Behavior and Energy Program. December.

TecMarket Works, Skumatz Economic Research Associates, Inc. (SERA), and Megdal and Associates. 2001. "Low Income Public Purpose Test (LIPPT) Report." For RRM Working Group Cost Effectiveness Committee.

Woolf, T., E. Malone, K. Takahashi, and W. Steinhurst. 2012. *Best Practices in Energy Efficiency Program Screening: How to Ensure the Value of Energy Efficiency is Properly Accounted For*. Synapse Energy Economics, Inc. pages 37 and 38. http://www.synapse-energy.com/Downloads/SynapseReport.2012-07.NHPC.EE-Program-Screening.12-040.pdf