

Policy and Other Intelligence from the Largest Non-Energy Benefits / Impacts (NEB/ NEI) Database and Model Around

Lisa Skumatz and Ann Gibbs, Skumatz Economic Research Associates, Inc.

ABSTRACT

One thing that comprehensive quantitative literature review gets you is data to support larger picture meta-analysis work. The authors assembled a database of non-energy benefits or non-energy impacts (NEBs / NEIs) values from a comprehensive literature review on the subject. This in-house database (called the “NEB-It” database) has more than 44,000 entries of NEB values, inputs, and associated program / measure / population features.

In response to requests from various utilities and state agencies, workshops, panels and conferences around the country, we have been applying information from the database to an array of policy issues. This paper presents results for a selected subset of these topics, including:

- Equity / hardship metrics, summarizing the literature and values to date for three alternative definitions of “hardship,” and the potential for particular NEBs to be used to reflect meaningful aspects of improvements in equity in response to utility initiatives.
- Assessment of State NEI adders, examining the state of the literature on the average values for estimated NEBs, and the prevailing opinions on which NEBs (within the perspective categories) have been considered priorities in different states.
- Methods Issues: Transferability of NEBs, and best practices based on a review of values from hundreds of US and international NEB studies.

The results from the previous work will be discussed, along with implications and potential next steps.

Introduction

Non-energy benefits / non-energy impacts (NEB/NEI) are the positive and negative impacts – beyond energy savings – that are provided as part of the delivery of the variety of energy efficiency programs and measures. These effects have been a research focus for several decades. Lists of NEBs had been postulated for some years, but were considered hard-to-measure (HTM). Initial measurement focused around low-income NEBs, because of their ability to help reflect progress in achieving policy goals for low-income programs (e.g., reduction in arrearages, improvements in comfort). Considerable work had also been conducted in quantifying emissions effects attributable to programs. Then broader work began that could chip away at the variety of impacts that had been omitted from the benefit-cost tests used for energy efficiency program screening. Instead of constructing these important benefit-cost ratios using all attributable benefits over all costs, HTM effects were missing. Including defensible estimates for these values could lead to modified prioritizations of measures or programs. This work has continued since the late 1990s, creating a large body of literature on NEBs.

As background, NEB values are realized by three groups of beneficiaries: program participants themselves (e.g., comfort), the utility (e.g., fewer bill-related calls) and society at large

(e.g., changes in emissions or jobs). Measurement of NEBs¹ began in the mid-1990s, and have experienced surges in research around 2001, 2007, 2014, and again, recently. Quantitative results on NEBs have been used in applications in program marketing / uptake, tracking progress toward policy/program goals (including low income), and improved utility benefit-cost tests for measure, program, and portfolio screening. program refinements, and cost-effectiveness tests. A number of literature reviews have been conducted that provide a flavor of the values to date (Skumatz 2009; NMR 2011; Itron 2016; Apprise 2018²), and other literature has reviewed or assessed quantitative methods applied (Skumatz 2002; Skumatz and Gardner 2005; Summit Blue 2006; Skumatz, Khawaja and Kropp 2009). Reports on the “state of NEBs”, gaps, and broad results and progress in adoption have also been prepared (NEEP 2017; Skumatz 2012-2016), and NEBs in cost-effectiveness tests, and state progress have been addressed in several studies (NEEP 2018; Skumatz and Malmgren 2014; Skumatz 2014-2018; NESP 2017).

Most NEB studies are “one-off”, estimating NEBs for one particular program or measure at one particular utility. This paper tries to illustrate the level of policy and other analyses that can be conducted when a comprehensive database is constructed, bringing together the quantitative information from more than 20 years of NEBs research and publications. SERA conducted this literature review and constructed a comprehensive, searchable NEBs (SERA’s “**NEB-It**”) database and associated SERA “**NEB-It**” model. The literature review of NEB studies included studies from the US and from 13 countries. The resulting assembled, searchable “NEB-It” database³ assembles all the quantitative NEBs and NEB inputs⁴ from this literature review. The database includes:

- 1800 NEB studies catalogued in the bibliography, 1200 with enough NEB content to review in detail.
- More than 700 studies with strong data, and 505 entered into the database.
- More than 44,000 lines of data of NEB-related values on 160 NEB categories.
- The database feeds inputs directly into the associated “NEB-It” NEBs model, developing tailored estimates of more than 130 NEB categories for up to 500 residential and commercial measures.

This paper uses this database to examine several policy topics to represent the types of higher-level questions that the *overall body* of NEBs research can answer – above and beyond what individual studies can provide for a specific program. For the analyses included in this paper, the study used subsets of the database’s NEB values relevant to the specific topic being analyzed. For example, we subsetting the “NEB-It” database entries on low-income programs and assembled averages of estimates of specific hardship NEBs for the research in Finding 3. We used averages and ranges of statewide NEB adders, and stacked the averages of all NEBs associated with each beneficiary perspective (participant, utility, and societal) for the results presented in Topic 4). In

¹ NEBs and NEIs represent the same concept, and both include positive and negative effects. The original terminology from the inventors of the concept was NEBs; NEIs is perhaps a better term to reconfirm positive and negative effects are included. However, for the remainder of this paper, we use “NEBs” to honor the inventors, to keep the literature under fewer monikers, and because it is possible to easily pronounce NEBs.

² And many others.

³ All data are searchable/sortable by NEB, measure, program type, sector (low income, C&I, etc.), year, units, state, primary vs. secondary data research, source, and many other NEB and program features.

⁴ Inputs include, for instance, “incidence” changes from programs, secondary data for valuing changes in incidence (both described later in the paper) and other input values like discount rates, lifetimes, etc.

general, we calculated averages and ranges, grouped NEB results by various categories by subgroups of populations or impact types, and developed totals and subtotals of NEBs so results could be compared and analyzed. This paper focuses on results related to: the volume of NEBs work and studies over time; hardship topics; health and safety findings; research on state NEB adders, and some methods-related suggestions.

Findings

Finding 1. The volume of NEBs research is increasing

The literature and database work involved review of hundreds of NEB studies. The database tracks the publication years of the studies. Figure 1 shows the number of NEB studies conducted by year, and the number with actual NEBs data by year. The figure shows:

- a significant share of studies addressing NEBs are papers without quantitative data, including NEBs policy papers, review of NEB concepts, or other documents.
- a significant number of studies have substantial data, including NEB results, or NEB inputs.
- In both cases, there are increasing reports over time.

NEBs research started to take off in the early 1990s, with significant increases in 2013 and a peak in the 2014-2015 timeframe. Note that some work in NEBs concepts (for example, emissions) was conducted very early on.

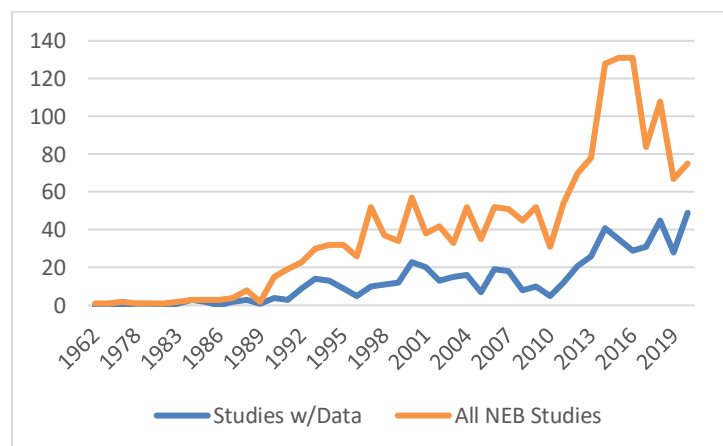


Figure 1: Growth in NEBs research: count of studies by year
Source: SERA NEB-It Database, Skumatz Economic Research Associates (SERA) Research 2021

Finding 2: There are four main quantification methods applied for NEBs

Detailed review of these hundreds of studies in the database shows that the NEB estimation methods used four main approaches used for calculating or developing estimates of individual NEBs categories.

- **Models:** Strong third-party models are available for estimating the economic output and job creation benefits from the manufacture, delivery, and installation of additional energy efficiency measures (IMPLAN™ and RIMS II™ are two of the most-used examples). The US has strong and very simple-to-use models for estimating the societal health effects from emissions reductions (EPA’s COBRA™ and AVERT™ models). These third-party models were the source of many estimates, and their use has grown over time. However, some utility or government studies (particularly international work, or early work) relied on local or tailored models for emissions, economic, or societal health effects
- **Change in incidence times secondary valuations:** These calculations are based on a change (delta) in occurrence that can be associated with a program or measure, times the value of that change. Examples include changes in specific illness occurrences times the cost of hospitalization or doctor visits or other appropriate valuations, changes in fire risk avoided from measures or interventions times the values in terms of property damage or injury or death, or as simple as the change in number of phone calls related to non-payment of bills times the minutes spent per call times the wage rate for utility staff making those calls. Many of the NEBs included in the database were calculated in this manner.
- **Financial calculations:** Financial calculations may overlap somewhat with incidence times value calculations, but examples include lower operations and maintenance (O&M) from installation of LED vs. CFL or other non-LED baseline bulbs using longer lifetime years times lower custodial staff costs per year, or similar calculations.
- **Survey valuations:** Some NEBs can only be derived from surveys. Comfort and noise might be able to be externally measured in some way (temperatures, draft/infiltration metrics, noise sensors) but that does not provide a measure of “comfort”, or noise that matters to participants. Studies conducted since 2000 have relied on sophisticated methods of valuing these effects through participant surveys (Skumatz 2000-2020; Summit Blue 2006; NMR 2011; and others), asking them to value these effects relative to the savings or through other methods. A large and increasing number of studies use these methods to measure comfort, internal and external noise, aesthetics, control over bills, and other NEBs with application to some benefit-cost perspectives, and to marketing applications.

Finding 3. Existing NEBs research can provide monetized feedback on at least 3 definitions of hardship impacts on vulnerable households – and the combination is worth 2-5 times energy savings

Many low-income weatherization programs include goals related to lowering hardship for participants. Unfortunately, the specifics of the “hardship” goal are almost never explicitly defined in program goals. The database allowed us to examine the range of NEBs that have been estimated for low-income programs, and select those most suited to three potential aspects of hardship:

- bill payment and financial safety,
- quality of life (QoL), and
- housing stock / housing retention effects.

These types of results can inform program and policy staff about whether existing programs have led to significant value to participating households in any of these alternate types of hardship. Using the database's average values for each of the NEBs assigned to these categories, the total NEB value from each of these definitions was calculated. These totals provide NEB values for those interested in any particular definition - whichever definition is most in line with the policy goals for the reader's state, utility, or program. The results are presented in Table 1.

Table 1: Estimated “Hardship” Benefits to Vulnerable / Low Income Households from Weatherization Programs, USD\$ per household per year.

Type of Hardship Metric	\$/household/year, basic hardship metrics	Multiple of average annual program energy savings – basic metrics	Multiplier of average annual program energy savings including enhanced definitions of hardship metrics
Bill payments & financial security	\$248	2.01	2.01
Quality of Life	\$215	1.74	-
Enhanced: Quality of Life including Comfort	-	-	2.20
Housing stock / housing retention	\$157	1.27	-
Enhanced: Housing stock including direct repairs*	-	-	1.59
Total All	\$620	5.02	5.80

Table note (*): Average \$240 of repairs annualized using 7-year life and 4% discount rate. Other assumptions easily substituted

Source: SERA NEB-It database, Skumatz Economic Research Associates (SERA) Research, 2021.

Calculation of the hardship values: The average for each NEB category assigned to the hardship metric was computed, and these individual NEBs were summed to develop the total dollar per household per year values shown in Table 1. We only include participant-perspective NEBs, as these hardship policies are participant-focused; utility and societal NBs are not included in these totals. We also excluded NEBs with few estimates in the literature, and those with inconsistent definitions in the published studies that led to wide variation in their NEB values.⁵ We omitted broad-ranging “health” estimates in favor of the total of NEBs associated with individual illnesses to avoid double-counting and overlaps in the totals. Finally, we omitted the value of home repairs (e.g., repair of broken windows, porch repairs, etc.) from programs that included these repairs. The annualized value of these benefits, which averaged \$240/participant from the data in the NEB-It database, should be added to the value of the housing stock hardship metric for programs that include these improvements. Values from studies using multiple valuations were translated to units of dollars per household per year prior to computing the averages of the component individual NEBs and calculating their totals.

Discussion of the Hardship Metrics and Results: The NEB totals and key NEBs included in each of the hardship definitions are discussed below.

⁵ Most particularly, we excluded fewer missed days from school. Some estimates of this impact included only the direct financial impacts (child care impacts), but others incorporated the longer-term financial impacts to the children missing school. To keep the estimates conservative and interpreted as near-term budget costs, we excluded this impact. For those interested in including the effect, its value averaged \$108/year/participating household.

- **Bill Payment and Financial Safety:** The NEBs leading to highest value to participant households included: fewer missed days from work (from fewer illnesses), avoided moving costs, reduced water / sewer / other bills, savings from reduced fire risk. Some commonly-estimated NEBs (arrearages, etc.) were substantially smaller value. The total value, based on averages in the “NEB-It” database is \$248/household/year, or about 2 times the value of the energy savings delivered by these programs.
- **Quality of Life:** The categories of NEB benefits included in the overall QoL valuation were limited to those representing illness effects, including fewer colds, asthma, and other illnesses, and reduced carbon monoxide risk / costs. We also constructed a broader definition that expanded the interpretation of the status of Quality of Life to include comfort-related effects to occupants. Adding comfort-related effects increased the average benefit value substantially. The total value of this hardship definition, based on averages in the “NEB-It” database, is \$215/household/year, or about 1.75 times the value of the energy savings delivered by these programs. Adding comfort benefits increase the multiplier to 2.29 times energy savings, as shown in Table 1.
- **Housing Stock / Housing Retention:** This largest value NEBs in this hardship definition included reductions in operational and maintenance costs (e.g., reduced repairs), and equipment lifetime extension effects, which can be suitably included for low income programs. The value for the sum of the components included in the of this hardship effect is \$157/household/year, or about 1.27 times energy savings, as shown in Table 1. Including an annualized value for house repairs, for programs including that element, increased the NEB multiplier value to 1.59 times energy savings.

Implications of the Hardship Results: These exploratory computations, which include the results of scores of low-income NEB studies, indicate that counting only energy savings would very substantially understate the benefits of these programs, and hardship effects from these programs are very sustainable to participating households. Furthermore, with the increasing attention on equity, a review of NEB values may provide a much more robust method reflecting progress from energy efficiency programs than do simpler metrics like “access” to programs. These metrics better reflect the value received by participants. Table 1 shows:

- by any of these definitions, the assistance to the vulnerable households participating in these programs is of greater value to the households than the direct energy bill savings delivered;
- the programs can take credit for providing 2-5 times the savings as a reduction in hardship, depending on whether one or all (non-overlapping) categories of hardship relate to the program’s policy goals (direct financial vs. broader definitions).
- these benefits are increased as definitions of policy are made more inclusive.
- direct financial benefits are especially large, but the quality of life and housing stock benefits also bring strong value to these vulnerable households.

Tailored estimates can, of course, be developed based on variations in program definitions / measures included, regions, policy goals and variations in hardship definitions, but these results indicate very substantial hardship benefits from programs to date.

Finding 4. Current state NEB “adders” are far short of values the literature could justify

Various states across the US use one or more of the five major benefit-cost tests (BCT) for screening measures, programs or portfolios. Economics would suggest BCTs should be constructed using the ratio of the present value (PV) of all attributable benefits over the PV of all attributable costs. However, historically, non-energy benefits attributable to energy efficiency programs were excluded from these computations because they were “hard to measure” (HTM). The omission of these NEBs leads to bias in these screening tests. To address this issue, at least partially, some states adopted “adders” as proxies for the missing NEBs, accepted as a “partial win” by stakeholders. This section addresses the question of whether the adopted adders appear to be sufficient proxies, based on the level of estimated NEBs in the literature.

The answer partly depends on which test the state uses. The common BCTs are the Utility Cost Test (UCT), Ratepayer Impact Test (RIM), Participant Cost Test (PCT), Total Resource Cost Test (TRC), and Societal Cost Test (SCT). These tests are ranked generally in the order of least broad to most broad perspective included, and thus, likely least extra NEB value missing to most omitted value. The state’s choice among BCT’s generally relates to its policy goals.

First, we inventoried the values of the adders that are currently in place. SERA’s review of the range and average values for state adders are shown in Table 2.

Table 2: Range and Average of State NEB Adders

	Electric Savings adders	Gas savings adders	Extra adders for Low income
Number of States	12	9	4
Range	10-30%	5-30%	10-15%
Average	13%	11%	12.5%
Mode	10%	10%	10%, 15%

In addition to the values in the Table, 16 states allow adders to be applied to commercial programs or measures, seven allow straight carbon-related (non-percentage) dollar per ton or dollar per kWh or similar adders, six states allow direct, easily- or reliably- measured NEBs, and three include extensive NEBs in their Technical Resource Manuals (TRMs) or similar reference documents.

Table 3: Comparison between Typical Adopted State NEB “Adders” and Average Estimated NEBs from the Database, by Perspective

	Typical Existing Adder in Place	Utility NEBs	Societal NEBs***#	Participant NEBs
Low Income	30%	25%*	200%	220%
Residential (excl. Low Income)	15%	15%	125%	160%
Commercial	15%	**	45%**	140%
Some or all of these values relevant for Specific Benefit Cost Tests:		UCT, TRC, SCT	SCT	PCT, TRC, SCT

Table Notes: (*) Excluding gas emergency NEBs and rate subsidy benefits, which can add more than 25%; (**) too few estimates for commercial programs found to report, or significant missing NEBs. (***) Omits regional economic/development adders – too few (normalized) observations; (#) excludes water system effects. Table includes conservative list of NEBs, rounded (excludes aesthetics, control/knowledge) and NEBs with few estimates. *Source:* SERA “NEB-It” database, Skumatz Economic Research Associates (SERA) Research 2021

SERA's review of the 20 states with adders finds twelve use the TRC as their primary test, four that use the UCT, four that use the SCT, and one that uses a state-specific test. These counts include one state that uses both the TRC and UCT.

The results in Table 3 indicate that a 15%-30% adder may not be highly biasing for the 4 states using the UCT, depending on what effects are elsewhere covered in avoided cost and other mechanisms. However, Table 3 shows that in no case do the percentage adders come close to covering the NEBs that could potentially be included, or even the most reliable or most local-policy-related portion (assume even one-half or one-third of the NEBs) that could contribute to each perspective. This omission is most biasing for the tests that include multiple perspectives (e.g., TRC or SCT), as the shortfalls for each perspective are compounded.

We recognize that the initial adders were adopted a decade or more ago, when the NEBs literature was less extensive or robust. Considering the current state of the literature, the apparent implications of these results are:

- The existing adders are substantially understating the omitted NEBs in the 12 states using the TRC and 4 using the SCT test.
- The current adders appear to dramatically understate the missing NEBs for low-income programs.
- The state with adders should revisit the adders and increase them to reflect at least the combined value for the highest-reliability NEBs, or should change their approach to allowing inclusion of specific, measured NEBs in the benefit-cost test.⁶
- The states that have not yet adopted NEBs should adopt considerably higher values than the 5-15% range that is most common.

The overall implication is that existing adders appear to be much lower than is likely justifiable from the existing literature, and that the adders used in existing states are a poor guide for "next" states looking to institute adders.

Finding 5: "Best Measurement Practices" and good reporting are important. Some NEBs tend to be over-estimated because studies are not careful about baselines.

The studies in the literature review that were identified as "methodological" were reviewed to identify some of the best practice for NEBs quantification. It is possible to identify some of the key "best measurement practices" (BMPs), and then compare the methods used in the various NEB studies to identify areas where BMPs have been applied less often.

Basic BMPs for NEBs. From early on, the literature (Skumatz 2000, 2002⁷) presented information on best measurement practices (BMPs) for NEBs. Core practices include;

1. Estimating "net" (and attributable) NEBs, meaning:
 - a) Net of Positive and negative NEBs,

⁶ Note that the State of Maryland has probably adopted the most progressive, but logical, language. In 2015, Order 87082 suggested that the test should include the NEBs that had been estimated in the most recent NEBs study, and going into the future, as additional NEBs estimates are developed that are for the relevant perspectives(s), they should be incorporated. See Malmgren and Skumatz (2014) and Skumatz (2014, 2019).

⁷ And others, now here augmented by other studies in the literature.

- b) Net of Net-to-gross (NTG) to avoid taking credit for NEBs not caused by the program.
- c) Net of standard efficiency equipment, meaning program-attributable NEBs should only take credit for the NEBs associated with the extra move to energy efficient equipment, net of any NEBs that would be delivered by installation of the alternative - baseline standard efficiency equipment.

However, our review of methods and values from the hundreds of sources examined for the database work provided an opportunity to observe the array of measurement and reporting practices in published reports. Resulting conclusions and recommendations from comparing strong studies to weak studies follow.

2. Our reporting and measurement recommendations for best practices include:
 - a) Clearly identify and present NEBs using three perspectives – utility, participant, and societal, using valuation methods appropriate to the beneficiary perspective. These three perspectives make it clear which NEBs are appropriate to include in the major Benefit-Cost Tests.
 - b) Use consistent units (usually dollars per participant per year or dollars per measure per year⁸), which can be added, summed, and most importantly, compared among NEBs to readily identify and compare relative sizes of NEBs. This metric supports computation of lifetime NEBs using weighted average measure lives (EULs).
 - c) Use appropriate discount rates for present value calculation based on perspective.
 - d) Where possible, estimate NEBs by measure to allow for the NEBs to be linked to the causal measures, and to provide NEB values that could be re-weighted as needed during evolution of program measure mix. This makes the NEB values more transferable across and between programs and to other jurisdictions. The most useful and transferable approach is to report the NEB value per participant receiving the measure, separate from the program-wide impact for that measure.
 - e) Where possible, examine NEBs by business type (for C&I) and by key residential customer groups (for residential measures and programs), in order to allow the NEB calculations to vary as programs may vary targets, or find and allow for revised NEBs estimates as participation patterns evolve over time.
 - f) Present NEB results in multiple terms to allow better transferability to the program over time, to be used by other programs, or to allow better comparisons between programs. For detail, see the section on transferability in a later “finding” in this paper.

Most of these BMPs are fairly straightforward to implement; however, a review of the database indicates that a number of these BMPs are not in place in all the literature. Deviations and deficiencies we found in the literature include:

- Units (2b): The most easily-identified deficiency is the lack of consistent units. Within single studies, NEB results are presented in annual terms for some, and present values for others, preventing comparison of large vs. small effects.

⁸ Some are measured percent of the energy savings per year per participant

- NTG (1b): Net-to-gross is often not addressed in the studies, and the uses to which these results are applied need to add NEBs up front with savings, or after-the-estimation, apply the NTG factor to estimate program-wide NEBs. Just like savings, when incorporating NEBs into BCT-related applications, the programs cannot take credit for NEBs that would have occurred without the program's influence.
- Measure-based NEBs (2d): These studies are still relatively uncommon, especially in the residential sector, but this is improving with more recent studies. Very few studies report the value for those receiving the measure, reporting only the program-wide value for the measure, which is dependent on an often-unreported percent of program participants that received the measure. That severely limits the ability to compare this value to programs with a different uptake in that measure or apply the value to the program over time.
- Baseline (1c): This is the most important error. Some O&M studies and many survey-based NEB studies do not take special care to note that the program-attributable NEBs must be only those associated with the difference from standard equipment that would be installed. More specifically, the baseline condition will vary based on whether the program is a Replace on Failure (ROF) program or "event", in which similar "standard efficiency equipment" or potentially industry standard practice (ISP) equipment may be the baseline. For early retirement, dual baselines may be needed, and there may be variations based on other events. That is essential for the benefit-cost work.
- Only the NEBs that make sense for the local policy goals and BCT perspectives are relevant for application to BCT. However, many studies total all NEBs as if all are equally relevant for all uses. Note that the full cohort of NEBs may be suitable for some applications, like marketing / advertising. Some studies have not linked back to goals, or sorted out which NEBs measured are useful for BCT vs. others that are really best suited to marketing applications.

The baseline deviation from best measurement practices is very important – and complicated. A review of the literature shows that many early (and recent) NEB studies gather information on the NEB value relative to the equipment being replaced. Instead, the correct methodology for program-related NEBs is to estimate the "extra" NEB impact delivered by the installation of an energy efficient program measure over and above that NEB impact that would have received from installation of a (new) standard efficiency model. While such hypotheticals can be difficult to answer, hypotheticals form parts of Net to gross and other EM&V surveys. The most complicated version to date appears to be a project the authors are directing that is estimating NEBs for residential heat pumps, which have multiple baselines, and multiple market "events" (replace on failure vs. early replacement), requiring multiple strata and a relatively large sample size. There are two reasons for making sure this convention is incorporated:

- Without it, NEBs would be overestimated – and NEB value above and beyond that attributable to the program would be estimated.
- Including this convention can provide NEBs for measure and market event combinations that allow for evolution of a program over time, and also improve the transferability of the NEB value to programs that are either replace on failure-based or early retirement, etc.

Finding 6: Measure-based NEBs studies are becoming growing in frequency, helping transferability. Review of the target audience is also necessary to improve transferability.

A key reason literature reviews of NEB studies are conducted is to identify whether there are NEBs that have been estimated in other programs that can be applied to a new program so that research money can be saved. This is transferability (Skumatz 2019⁹). Of course, NEBs estimated for a specific program are preferred, but transferability, and the chance to compare the consistency of results also has value. The use of the data in the database is substantially enhanced with data that can be considered transferable.

Measure-based analysis enhances transferability of NEBs. However, the literature review conducted for the database shows that in the early phases of NEBs research, most NEBs in the residential sector were estimated on a program-wide basis. Commercial NEBs were estimated at the measure level or the program level, depending on the type of program, e.g., estimates of air-compressor or motors programs were estimated at the measure level, but “new construction” programs were commonly measured at the program level. Based on our experience trying to apply NEBs to programs, it is preferable to develop NEBs at the measure level because:

- The NEB and its cause are directly linked, and the NEBs can then be expressed as by NEB category, and NEBs per measure installed.
- This supports calculation of varying NEB levels when the number or mix of specific measures evolves over time within a program.
- This also potentially supports transferability of the NEBs across programs, and potentially allows use of some NEBs across jurisdictions, depending on the similarity of climate (for climate-affected measure and NEBs), and other factors that affect transferability of NEBs (Skumatz 2019¹⁰).

The best approach is to stratify NEBs surveys based on measures, with question batteries that address associated NEB effects and relevant changes in incidence of NEB impacts associated with that measure. Note that for C&I NEBs, there is likely a case to consider stratifying NEB surveys by business sector as well. A number of non-survey-based NEBs (e.g., maintenance, deferred replacement), financial or other calculations are naturally conducted on a measure basis. The literature review conducted for the database development finds there have been several attempts to assigned portions off program-wide NEBs to particular measures after-the-fact. Our review found two approaches have been used:

- Some early studies allocated these program-wide NEBs to individual causal measures (Skumatz 2018¹¹) using the simplistic approach of sharing to measures using measure savings. This method has a number of key weaknesses, the most important of which is not allocating NEBs to clearly causal measures that yielded no savings or negative savings (comfort from fixing a non-functional HVAC system), and other practical

⁹ A detailed discussion of which NEBs can be transferred between different studies is included in Skumatz 2019. This addresses NEBs that are measure-independent (emissions, etc.), location-dependent (economics, comfort, and others), and so on. The study also discussed dependent factors, possible adjustment methods and other topics related to transferability.

¹⁰ See Skumatz (2019) IEPEC for discussion of the issue of factors affecting transferability.

¹¹ See Skumatz (2018) ECEEE for discussion of this topic

difficulties. Improvements on this simple method include: using energy costs as an allocation method, or using an estimated relationship between savings and energy cost, and assigning the “fitted value” when energy savings are zero or missing.

- A more refined approach to the issue of using program-wide information to allocate NEBs to measures when measure-based information was not available used regressions. This approach was demonstrated in the residential sector in Skumatz 2006, and regressed NEBs for each respondent against the measures they did and did not have. These results were used in a number of NEB models. However, measure-based information is the most flexible, allowing program NEB estimates to vary as measure mix may change going forward.

Understanding demographic (or firmographic) differences between program participant is also important to improve transferability of resulting NEBs. Measure-based results are not enough to assure strong transferability. Program targets population or program participant differences also affect participant valuations of NEBs for the same measure. Households with chronically ill residents, or with elderly residents or young children / home-all-day households, may report different NEB values for comfort and other important NEBs and using those values for programs with a dramatically different participant mix will not be appropriate. A review of the database shows that this has almost never been addressed in residential studies. To incorporate this effect would require potentially pre- or post-stratifying on some of these factors, some of which have been statistically demonstrated to be significant (Skumatz 2006, 2019). On the commercial side, this may suggest stratifying by business type¹² or other factors to allow for application of the NEB estimates to changes in program participation mix in the future.

Measure-based stratification improves transferability, but considering comparability of participant mix is also important.

Finding 7: The findings from NEBs literature could be better leveraged if studies reported the values more consistently and thoroughly.

As mentioned, transferability of NEBs from one study or one program to another can save money and build confidence through comparability of results. However, this objective has been hampered to date by poor reporting-out procedures by NEB studies.

Normalization information. Commercial NEBs could and should be expressed with normalization factors, and that could resolve a good degree of the “size” disparities for NEB studies. A key element is providing NEBs by measures (not by program). However, the very large variation in commercial NEBs results by measure are likely explained by the size of the business into which the measure was installed.

A standard for residential NEBs is reporting the value in dollars per participant (or unit or household) /year. A review of the NEB-It database shows commercial NEBs are expressed in units that vary greatly from study to study, even when the same author is present, but the most common value is per business or program-wide. These particular values make it difficult to compare between studies, and offer little value to other programs with different sized participants.

¹² Several studies have done this, including Coates et.al. (2000), and DNV-GL (2018), among others.

To improve transferability of NEB values for at least commercial NEBs, some normalizing factor should be associated, in order to allow the numbers to make sense as program participant mix or other elements evolve. Two attractive options include:

- Size indicator of square footage (or if that is not obtainable, perhaps number of employees)
- Or translating and expressing the NEB as a percentage multiplier relative to savings for key NEBs (as is done in the residential sector).

A number of C&I studies have presented NEBs as a percentage multiplier or per square foot option.¹³ Some of these results only have one or two values (only one or two studies), but the approach shows promise for improving transferability of NEBs as a program evolves or to other programs.

If NEB studies expressed the NEB value *also* in terms of per kWh (or possibly per employee, per square foot, per dollar spent, or similar normalizing metrics) comparability and transferability would be improved. Especially important is presenting background information on the program (e.g., the energy savings for participations), and separate reporting of the NEB values for those receiving the measures, as well as the overall program-wide results weighted by the percentage mix of measures delivered by the program. This allows results to be transferred across programs, and allows the result to be used as measure mix evolved over time within a program. Our database review found that the minority of studies had provided information using these formats.

Report the value for those with the measure. Report the NEB result both for those experiencing the effect or getting the measure, not just the program-wide results. Then the value that is reported in the literature can be compared because it will be absent the effect of the measure mix in the specific program. This is short, but vitally important; these separate values are very rarely reported.

Conclusions and Recommendations

Twenty years of research monetizing values for participant, societal, and utility NEB effects provide useful information for: marketing/uptake; tracking progress toward policy/program goals (including low income); and improved, less biased, utility benefit-cost tests for measure, program, and portfolio screening.

For the majority of the analyses presented in this paper, we used SERA's "**NEB-It**" database, containing information from a comprehensive review of 20-plus years of literature (1900 studies), and more than 44,000 lines of NEB data from more than 500 NEB studies covering 160 NEBs. Combined with the associated NEBs model (SERA's "NEB-It" model), that derives tailored estimates of more than 130 NEBs excluding Tier 3 ("not well-estimated") categories (Skumatz 2019), the authors were able to examine specific program, policy, research, and methods questions, as well as the NEB gaps and next issues also highlighted in this paper. The paper reports on seven findings.

¹³ including work by the authors for NYSEERDA (Skumatz et al. 2005), and work by the author estimating NEBs for a commercial retro-commissioning program presented the NEB results on a per-square foot basis (Skumatz 2004), and other studies using these normalizing factors.

- Finding 1: The volume of NEBs research is growing, and the literature is increasingly robust in number, topics, sophistication, transferability, and coverage.
- Finding 2: There are four main NEBs estimation methods used in studies: models, incidence times secondary value, financial calculations, and survey-based estimates. Each has been well-applied to a variety of the NEBs in the literature.
- Finding 3: Using any of three types of “Hardship” definitions, hardship benefits significantly exceed the value of energy savings, and the combination is five times as valuable to program participants as the energy savings. Significant hardship benefit is delivered by these program efforts. However, clarity on the specific aspects of hardship that a policymaker is targeting for the program would help NEB studies direct their measurement efforts to those NEBs that would reflect progress in achieving goals.
- Finding 4: Current State NEB Adders are far short of values the literature could support. Many adders are a decade old, or are adopted from states whose adders are a decade old. The literature has advanced significantly from that time, and the analysis of the current literature indicates proxy adders for states that use the TRC or SCT are substantially lower than would likely be assigned if up-to-date literature was examined.
- Finding 5: Best Practices and good reporting out are important. Some NEBs tend to be over-estimated because they are not careful about baselines. Important best practices include presenting results in consistent units, clearly assigning NEBs to beneficiary category, reporting NEBs for those receiving the measure as well as weighted by the share of program participants receiving the measure, and especially taking care that the NEB included in BCT are triple-net: net of positive and negative impacts, net of net-to-gross, and net relative to the proper baseline.
- Finding 6: Measure-based NEBs are more useful because of their transferability, and are growing in frequency. This allows much more flexibility in using the NEB results as programs vary measure mix in the future. Consideration of variability of NEBs based on characteristics of participants would also enhance transferability going forward.
- Finding 7: The findings from NEBs literature could be better leveraged if studies reported the values using more consistent units and provided supporting program, measure, participation, and savings information that allowed the result to be normalized, scaled, and more easily transferred and updated.

References

- Apprise, 2018. *Connecticut Non-Energy Impacts Literature Review: R1709 Final Report*. Prepared for Connecticut Energy Efficiency Board (CTEEB). CT. December.
- BeMent, D., and L. Skumatz, Ph.D., 2007. “New Non-Energy Benefits (NEBs) Results in the Commercial / Industrial Sectors: Findings from Incentive, Retrofit, and Technical Assistance / New Construction Programs.” France: *Proceedings from the European Council for an Energy Efficient Economy (ECEEE)*.
- Coates, B., D. Pearson, and L. Skumatz, 2000. *Operations Resource Assessment Service: Process and Impact Evaluation*. Prepared by Seattle City Light Evaluation Unit, Energy Management Services Division, and Skumatz Economic Research Associates. Seattle WA. May 2000.

- DNV-GL, 2016. *Stage 2 Results – Commercial and Industrial New Construction Non-Energy Impacts Study – Final Report*. Prepared for the Massachusetts Electric and Gas Program Administrators, March 24.
- Imbierowicz, K., L. Skumatz, and J. Gardner, 2006. "Net NEB Multipliers for Economic Impacts: Detailed Analysis of Differences by Program Type and State." Asilomar, CA: *Proceedings of the ACEEE Summer Study on Building Conference*.
- Itron, 2014. *Development and Application of Select Non-Energy Benefits for the EmPOWER Maryland Energy Efficiency Programs*.
- Magouirk, J. 1995. "Evaluation of Non-Energy Benefits from the Energy Savings Partners Program". Chicago: *Proceedings from the International Energy Program Evaluation Conference (IEPEC)*.
- Malmgren, I., and L. Skumatz, 2014. "Lessons from the Field: Practical Applications for Incorporating Non-Energy Benefits into Cost-Effectiveness Screening." Asilomar, CA: *Proceedings of the ACEEE Summer Study on Buildings*.
- National Efficiency Standard Project. 2017. *National Standard Practice Manual*.
- NEEP. 2017. *Non-Energy Impacts Approaches and Values: An Examination of the Northeast, Mid-Atlantic, and Beyond*.
- NMR Group. 2011. *Massachusetts Special and Cross-Sector Studies Area, Residential, and Low-Income Non-Energy Impacts (NEI) Evaluation*.
- Skumatz, L. 1997. "Recognizing All Program Benefits: Estimating the Non-Energy Benefits of PG&E's Venture Partners Pilot Program (VPP)." Chicago: *Proceedings of the 1997 IEPEC Conference*.
- .2002. "Comparing Participant Valuation Results using Three Advanced Survey Measurement Techniques: New Non-Energy Benefits Computations of Participant Value." Asilomar, CA: *Proceedings of the 2002 ACEEE Summer Study on Building Conference*.
- . 2006. "Methods for Measuring Non-Energy Benefits and Attributing Program Effects." London: *Proceedings of the 2006 EEDAL Conference*. June.
- . 2014. *Non-Energy Benefits / Non-Energy Impacts (NEBs/NEIs) and their Role and Values in Cost-Effectiveness Tests: State of Maryland*. Prepared for Natural Resources Defense Council (NRDC). New York. March.
- . 2015, "Efficiency Programs' Non-Energy Benefits: How States are Finally Making Progress in Reducing Bias in Cost-Effectiveness Tests." *The Electricity Journal*, September.

- . 2016. “Risks associated with NEBs.” Rome, Italy: *Proceedings from the IEPPEC Conference*.
- . 2017. “Soup to Nuts on NEBs – Methods, Results, and Application at the Utility and the Regulatory Level.” Irvine, CA: *Proceedings of the EEDAL Conference*.
- . 2019. “Non-Energy Impacts (NEBs/NEIs) Beyond Literature Review – New Findings on Values, Updated Models/Tools, and Suggestions for Expanded Use Across the US.” France: *Proceedings of the European Council for and Energy Efficient Economy (ECEEE)*.
- Skumatz, L., C. Dickerson. 1997. “Recognizing All Program Benefits: Estimating the Non-Energy Benefits of PG&E’s Venture Partner Pilot Program (VPP).” Chicago: *1997 Energy Evaluation Conference*. August.
- Skumatz, L., C. Dickerson, and B. Coates. 2000. “Non-Energy Benefits in the Residential and Non-Residential Sector: Innovative Measurements and Results for Participant Benefits.” Asilomar, CA: *Proceedings of the ACEEE Summer Study on Buildings*. August.
- Skumatz, L., and J. Gardner. 2005. “Methods and Results for Measuring Non-Energy Benefits in the Commercial and Industrial Sectors.” *Proceedings from the IEPEC Conference*. August.
- Skumatz, L., M. S. Khawaja, and J. Colby. 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*. Prepared for CIEE Behavior and Energy Program, CIEE, California Public Utilities Commission. Berkeley, CA. November.
- Skumatz, L., M. Santulli, and D. D’Souza. 2019. “Meta Results in NEBs / NEIs – Progress in NEB Values, Attribution to Measures, and State Adoption into Cost-Effectiveness Tests.” France: *Proceedings from 2019 ECEEE Conference*.
- Smith-McClain, L., L. Skumatz, and J. Gardner, 2006. “Attributing NEB Values to Specific Measures: Decomposition Results from Programs with Multiple Measures.” Asilomar, CA: *Proceedings of the ACEEE Summer Study in Buildings*.
- Summit Blue Consulting LLC and Quantec, LLC. 2006. *Non-Energy Impacts (NEI) Evaluation*. Prepared for New York State Energy Research and Development Authority. June.
- Weinsziehr, T., and L. Skumatz. 2016. “Evidence for Multiple Benefits or NEBs: Review on Progress and Gaps from the IEA Data and Measurement Subcommittee.” Amsterdam: *Proceedings of the IEPPEC Conference*. June.