SKUMATZ ECONOMIC RESEARCH ASSOCIATES, INC.
■ Consulting to Government & Utilities

Boulder Office: 762 Eldorado Drive, Superior, CO 80027
Voice: 303/494-1178 FAX: 303/494-1177
email: skumatz @ serainc.com
Website: www. serainc.com; payt.org

SERA
SKUMATZ ECONOMIC RESEARCH ASSOCIATES

NON-ENERGY BENEFITS / NON-ENERGY IMPACTS (NEBs/NEIs) AND THEIR ROLE & VALUES IN COST-EFFECTIVENESS TESTS: STATE OF MARYLAND

FINAL REPORT

Prepared by:
Lisa A. Skumatz, Ph.D.
Skumatz Economic Research Associates, Inc. (SERA)
762 Eldorado Drive, Superior, CO 80027
303/494-1178 skumatz@serainc.com

Prepared for:
The Natural Resources Defense Council, Inc. (NRDC)
40 West 20th Street, New York, New York 10011
Contact: Deron Lovaas

Policy White Paper / Expert Report, FINAL REPORT
March 31, 2014

Contents

*Citation: 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 Natural Resources Defense Council (NRDC), Superior, CO, March.

Acknowledgements: Many thanks to Dana D'Souza and Dawn BeMent, SERA staff, for valuable contributions to this research paper.

About the Author: Dr. Skumatz, an economist / econometrician, Principal of Skumatz Economic Research Associates (SERA) and founder of the non-profit Econservation Institute. She has 34 years of experience in energy efficiency program and policy evaluation. Lisa has specialized in process and attribution evaluation and market assessment work, and has conducted pioneering work in both measure lifetimes and non-energy benefits (NEBs) and cost-effectiveness, the focus of this paper. For two decades, she has been a national leader in NEBs, with more than 50 peer-reviewed publications on this topic alone, and scores of US and international conference presentations. Her NEBs work has been cited in state deliberations on cost-effectiveness tests across North America.

0. ABSTRACT

States across the nation are reconsidering the definition and design of cost-effectiveness test procedures used in the energy efficiency regulatory arena, including the State of Maryland. Twenty years of research and measurement of traditionally-omitted program impacts, or non-energy benefits (NEBs), have provided increasingly robust and consistent results. The regulatory tests are designed to assess costs and benefits, but protocols omitted some benefits, presumably because reliable values were not available. This leads to computational bias in benefit-cost ratios (from the omission of net benefit categories, but not omission of costs), and as a result, bias in decision-making using these ratios. Zero is the wrong proxy value; research has proceeded, and the results for a number of subcategories of NEBs can be properly reintroduced into these regulatory tests. Revising the tests (TRC, Societal Tests, or whichever others best reflect the state's energy goals) and incorporating subsets of NEBs reduce sources of bias in program and portfolio decision-making, and more appropriately directs the investment of millions of public or shareholder dollars.

The literature on NEBs has evolved through several levels of maturation,[1] and now consists of more than 300 studies of various types including results from programs around the country. Modeling methods[2] and consistency of values for many utility-, societal-, and participant-perspective NEBs has improved, and NEBs have become more familiar, through their use in marketing and other applications across North America. As a consequence, about a dozen states have already come to include some NEBs-related treatment in their regulatory benefit-cost testing procedures. Iowa, Colorado, Oregon, Washington, Vermont, New York, DC, and others include at least simple adders (between 7.5% and 25%) reflecting subsets of NEB contributions. Massachusetts, Vermont, Colorado, New Hampshire, BCHydro, Oregon, Connecticut, Rhode Island, Maine, DC, and others allow inclusion of subsets of "readily measured" or specific NEBs in benefit-cost tests, and the list is growing.

Twenty years of research on NEBs has developed more robust estimates; however, incorporation into benefit-cost tests lagged for three reasons:

- Chicken and egg issues. High quality values lagged because there was very limited funding of NEBs estimation work, since they weren't incorporated into use in applications with real value to the utilities or regulators. Consequently, use in tests lagged because there were concerns about the quality of the values.
- There isn't yet agreement on quality values for all NEBs categories, and there was concern that the estimates of some NEBs might not be accurate enough. In the near term, *inclusion of some* NEBs is better than *exclusion* of NEBs; each value helps reduce bias in tests. In addition, even if

---

[1] The evolution included four stages. Stage 1 (1994-1998) involved background organizing NEBs into perspectives, identify measurement principles for "net" NEBs, and preliminary estimations of two dozen categories. Stage 2 (1998-2001) included early rounds of documented derivations / estimates of NEBs, suggested incorporation into B/C tests, refinement of three main NEB estimation methods (models, incidence times valuation, and survey-derived estimates), and work on academic basis for survey approaches. Stage 3 (2001-present) included continuing expansion of estimates to more types of programs, enhancements of best practices, increasing familiarity of NEBs among stakeholders, application to marketing, and peer reviewed publications of results. Stage 4 (2008 to present) includes a period of refocus on the role of NEBs in regulatory and benefit-cost test applications.

[2] Estimation methods representing state of the art fall into three main approaches. Modeling approaches like third-party input-output models are used to estimate net economic multipliers from transferring dollars from generation to industries affected by energy efficiency programs, and models are also used for emissions impacts, and potentially reliability, etc. 'Incidence times value' approaches use primary and secondary data to estimate the value of program-related changes, for example arrearage studies, value of fewer bill-related calls or fewer emergency incidents, etc. Finally, very specialized comparative / ranking survey-related approaches (with basis in academic literature) are used to determine values of several important participant-side NEBs, including comfort, value of reliability, etc.

a precise point estimate isn't available, if the high and low ranges of a NEB don't change the program decision, the information is improved over using a zero value (perfect as enemy of the good issues).

- NEBs are perceived as expensive to estimate.[3]

These concerns led to continued use of "zero" as the value used for a variety of omitted[4] benefits, rather than computing benefit-cost ratios using parallel treatment of benefits and costs. Including NEBs – even subsets of NEBs – would reduce bias in the billions of dollars that are invested in energy efficiency programs across the nation.

An inventory of state regulatory procedures has shown that this landscape has changed. We see incremental progress – but clear and distinct progress – toward addressing the bias inherent in tests that exclude NEBs. A domino effect has resulted; as one state makes progress, another directly incorporates that progress into their next round of deliberations. Improved values for Maryland can be used to help improve the allocation of funds among energy efficiency programs – and between generation alternatives and energy efficiency.

This SERA study conducted a review of the "state of NEBs", especially as related to residential weatherization programs. The study reviewed the values – in dollar and percentage terms – estimated for NEBs from a large number of weatherization programs, and provides summaries of the ranges and typical values for the NEB categories.[5] The relative size of NEB values (in percentage and dollar terms), the consistency of their estimated values from quantitative studies, and the degree to which the NEB category will vary or not vary with program types is reflected in the columns in Figure 0.2. Consistent estimates are available in about twelve major NEBs categories and about 20 individual categories (see Figure 1.1). Several main categories are relatively invariant with program type and vary fairly directly with the amount of avoided generation (emissions, which may vary some with peak / off-peak generation mix, and T&D loss / reliability / energy infrastructure factors[6]). Most of the other NEB categories vary with the program's measures (comfort mostly results for those programs affecting shell or HVAC measures), or based on target audiences (low income programs bring additional NEBs).

Based on the analysis, short-term, medium-term, and longer-term recommendations for a NEBs strategy for Maryland is provided below. The short-term recommendation in Figure 0.1 incorporates values from a subset of NEB categories that have consistent results, and which are appropriate for inclusion in the TRC. None of these recommendations incorporate the full values for estimated NEBs; a conservative

---

[3] Considering for value-based decision-making, this may not be true. Many of the most important NEBs can be incorporated into existing process evaluations with marginal cost increases. Arrearage studies are already conducted. Comparing the "bang for the buck" for possible improvements in the overall accuracy of benefit-cost tests, another impact evaluation on a mature or little-changing program might change the benefits (savings) estimate a few percent. Deferring an impact evaluation and conducting a NEBs study would lead to benefits estimate improvements and reduction of bias many times that amount, based on the "math" of a B/C test.

[4] NEBs were often called "hard to measure" (HTM) effects.

[5] The table includes ranges and "typical values" for major NEBs categories. "Typical" values were defensible values selected based on a review of mean, median, and clustering of results from multiple studies. A total of more than 20 studies were reviewed, including studies in the Northeast / Mid-Atlantic region. Some studies provided only dollar values; others provided values in terms of multiples of bill savings; other provided both. There were also variations in terms of which NEB categories were and were not included in the estimation work. Where possible, we translated to consistent units. However, the studies available for the two comparisons (percentage multiplier relative to bill savings, vs. dollar adders) were not the same. For this reason, the ranges and "typical" values will not quite translate between the two treatments. Percentage adders may be the simpler treatment, allowing computation of a multiplicative adder onto existing bill savings in the B/C ratio computation, and allows simpler translations to scaling of program sizes. However, both provide valuable information.

[6] Note that T&D, line loss, and environmental compliance values (or some subsets) are already included in avoided cost figures for energy for some utilities, and should not be double-counted in those cases.

approach was taken for the short term, incorporating less than half, or a fifth, of the total of typical values from categories that are regularly estimated. We recommend Maryland implement a hybrid approach to the inclusion of NEBs. Factors that relate directly to savings and avoided generation can be a constant value for all programs (emissions, T&D / reliability / infrastructure)[7]; program-specific adders can then be "swapped out" based on the program type and targets.

The values presented for Maryland's weatherization program are the total of the following NEB categories:

- All NEBs: includes valued impacts from utility arrearages, environmental GHG emissions, participant comfort / noise, participant health/safety impacts, along with several "to consider adding" including valued impacts from economic multiplier effects; participant home value / improvement impacts, and participant water bill savings.
- Excluding NEBs "to consider adding", specifically valued impacts from economic multiplier effects; participant home value / improvement impacts, and participant water bill savings.
- The difference between "somewhat" conservative vs. "very" conservative values are the inclusion of estimates from improvements in home appearance, and from reductions in low income subsidies from reduced energy usage (a benefit to all ratepayers).

The emissions elements could be applied for any program; the other values are program-specific.

The study developed mid-term and longer-term recommendations; each phase includes improved NEB estimates. The mid-term recommendations include conducting several low-cost, fast-turnaround studies to develop Maryland-tailored values; suggestions include a participant-side survey, a Maryland-based economic multiplier (considering Maryland industries affected), work on a somewhat-refined emissions computation, and values from the multipliers table for other values. In the longer-term, we recommend the mid-term recommendations plus incorporating participant-NEBs surveys into occasional, periodic process or impact evaluation studies, periodic arrearage studies, and updated literature values.

Figure 0.1 Summary of Recommended NEB Value Adders for Maryland Weatherization (Wx) Program

	Somewhat conservative	Very conservative
Total All (recommended and "to consider" adding)	Base (Emissions): 12% adder (or 1.7 cents/kWh for Maryland) (7% from literature) Plus Wx-specific adder:[8] 80% (or $124) plus economic multiplier 0.69 times program expenditures per household (or $60) Plus Low income adder: 16% ($13) if low income subsidies in place	Base (Emissions): 12% adder (or 1.7 cents/kWh for MD) (7% from literature) Plus Wx-specific adder: 55% ($82) plus economic multiplier 0.69 times program expenditures per HH (or $60) Plus Low income adder: 16% ($13) if low income subsidies in place
Total excluding those that should be "considered	Base (Emissions): 12% adder (or 1.7 cents/kWh for MD) (7% from literature) Plus Wx-specific adder: 41% adder (or $73)	Base (Emissions): 12% adder (or 1.7 cents/kWh for MD) (7% from literature) Plus Wx- adder: 35% adder ($49)

Table Note: Percentage items are used by adding the percentage to the energy savings value in the B/C test. The value in dollar terms would be incorporated by adding $x per household (per year) in net benefits attributable from the program.

[7] *We did not find extensive literature on the reliability factor, although this would be an appropriate addition to this base factor for the hybrid adder. Factors for T&D losses have been estimated and applied in a few locations. In some utilities, T&D, line loss, and environmental compliance values (or some subsets) are already included in avoided cost figures for energy, and should not be double-counted in those cases.*

[8] *The included NEB values used follow. Utility carrying cost on arrearages (2%, $2.50-$4; higher for low income). Utility avoided low income subsidy, if offered by the utilities (16% adder or $13). Societal economic multiplier impacts (Conservative multiplier from a weatherization study is 0.69 applied to (non-administrative) dollars spent for program; other values from literature are adders of 31% / $60). Participant comfort / noise / light adder is about 10% / $18 for comfort alone; adding the other impacts is a total of 26% / $69). Participant health and safety impacts are about 13% / $16.50. Home improvement impacts NEB values are about 18.8% / $36; excluding aesthetics, the values are 10% / $18. Participant water bill savings impacts are 20% or $15 from other studies.*

Figure 0.2 Summary of Ranges and "Typical" Values for NEBs for Weatherization Programs

Note: Relative consistency indicator: ** low variation / relative consistency across programs; * low variation / relative consistency within program types; ~somewhat consistent; Variations by program, target audience, or limited variation by program are noted in the last column.

Subtotals by major categories Weatherization Programs	Dollar NEB Values Range Low-High		Typical Value	Percentage NEB Values Range Low-High		Typical Value	Consis-tency	Varies with Pgm Target Audience,
UTILITY PERSPECTIVE								
Payment-related	$2.55 -	$14.50	$6.40	1% -	14.5%	4.7%	*	Pgm
Added if Low Income subsidies avoided	$3.00 -	$25.00	$13.00	4% -	29.0%	16.4%	*	Pgm & target
Service Related	$0.10 -	$8.50	$3.25	0.1% -	2.7%	0.8%	*	Pgm
Other Primary Utility	$0.13 -	$2.60	$1.40	2.1% -	3.3%	2.4%		
TOTAL UTILITY NEBs	$5.78 -	$50.60	$24.05	7.4% -	49.5%	24.4%		
UTILITY NEBs MULTIPLIER	3% -	25%	12%	0.4% -	14.8%	3.3%		
SOCIETAL PERSPECTIVE								
Economic	$8.00 -	$340.00	$115.00	3.0% -	237.6%	31.1%	*	Pgm
Environmental / Emissions	$3.00 -	$180.00	$60.00	0.7% -	57.9%	7.1%	**	Ltd variation
H&S equipment / fires	$0.00 -	$0.30	$0.00	0.3% -	0.3%	0.0%		Pgm
Health Care	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%		Pgm
Water / Wastewater infrastructure	$1.00 -	$28.00	$15.00	0.9% -	33.1%	17.0%		Pgm
TOTAL SOCIETAL NEBs	$12.00 -	$548.30	$190.00	5.0% -	329.0%	55.3%		
SOCIETAL NEBs MULTIPLIER	6% -	274%	95%	4.4% -	295.5%	36.5%		
PARTICIPANT PERSPECTIVE								
Water and Other bills	$2.85 -	$54.00	$15.00	4.5% -	63.4%	20.0%	*	Pgm
Financial / customer service	$0.27 -	$36.70	$3.60	8.7% -	16.4%	3.4%	*	Pgm & target
Economic Dev'p / Hardship	$0.00 -	$115.00	$75.00	26.3% -	55.3%	8.0%		Pgm & target
Equipment Operations	$26.00 -	$127.00	$82.00	17.1% -	42.7%	28.4%		Pgm
Comfort, Noise, Related	$26.00 -	$105.00	$69.00	12.2% -	51.3%	26.6%	*	Pgm
Health / Safety	$3.02 -	$100.50	$16.50	1.5% -	59.5%	12.8%	*	Pgm
Control / Education and Contributions	$26.25 -	$177.00	$89.75	19.8% -	72.0%	26.2%	*	Pgm
Home Improvements	$10.50 -	$77.00	$36.00	8.3% -	38.4%	18.8%	~	Pgm
Special / reliability / other	$0.00 -	$4.05	$0.00	0.0% -	4.8%	0.0%		Ltd, target
TOTAL PARTICIPANT NEBs	$94.89 -	$796.25	$386.85	98.5% -	403.8%	144.1%		
PARTICIPANT NEBs MULTIPLIER	47% -	398%	193%	98.5% -	403.8%	144.1%		
All NEBs Multipliers: Relative to Bill Savings								
Utillity	3% -	25%	12%	0% -	15%	3%		
Societal	6% -	274%	95%	4% -	296%	37%		
Participant	47% -	398%	193%	99% -	404%	144%		
ALL Multiplers - relative to bill savings	56% -	698%	300%	103% -	714%	184%		

NOTE: Ltd variation for emissions are for peak / off-peak focused programs.

1. EXECUTIVE SUMMARY AND RECOMMENDATIONS

1.1 Introduction

States across the nation are reconsidering the definition and design of cost-effectiveness test procedures used in the energy efficiency regulatory arena, including the State of Maryland. Twenty years of research and measurement of traditionally-omitted program impacts, or non-energy benefits (NEBs), have provided increasingly robust and consistent results – results that are suitable for regulators, utilities, and interveners to explore options for updating cost-effectiveness test methodologies to better inform and address potential biases in program and portfolio decision-making. The literature on NEBs now consists of more than 300 studies of various types, and more than a dozen states currently include some NEBs-related treatment in their regulatory benefit-cost testing procedures.

1.2 NEBs: Progress in Research and Measurement

NEBs represent the array of positive AND negative impacts provided by energy efficiency programs beyond kilowatt-hour savings. They accrue to three classes of "beneficiaries":

- utilities and their ratepayers (largely changes in costs and revenue payment patterns);
- society (changes in emissions, job creation / economic development, water infrastructure, broad health effects, neighborhood impacts, etc.); and
- participants (ranges of impacts to residential and commercial participants including changes in comfort, maintenance, productivity, health, indoor air quality, etc.).

Figure 1.1 provides a list of common NEBs, grouped into the three perspectives.

Progress in Four Stages

Twenty years of research have included four main stages of research, which are summarized below. More specific detail on the research progress made during these stages is provided in Chapter 2.

- Stage 1 (1994-1998): In this phase, the research established beneficiary categories and basic measurement lessons; developed and assessed measurement methods and strength / weaknesses; pioneered approaches for hard-to-measure participant NEBs; and developed basic best practices for attribution of NEBs.

- Stage 2 (1998-2001): This stage included the first rounds of documented derivations / estimates in multiple NEBs categories; identified benefit/cost tests as an appropriate application of NEBs, and explored the refinements necessary for definitions of enhanced tests; and resulted in research projects that estimated more than three dozen categories of NEBs for programs in a variety of sectors and program types. Research during this stage began to be published, and work established and solidified three main measurement approaches (discussed below). Studies conducted initial comparisons of results and approaches between programs; and concluded with a detailed / documented study of three dozen low income NEBs reviewing more than 200 sources and options.

- Stage 3 (2001-present): In this stage, the NEBs estimation work expanded to many programs for utilities around the nation, resulting in numerous publications at peer-reviewed conferences in the US and Europe. Best practices approaches and NEBs issues (measure-based NEBs, new measurement methods; comparisons between results; volatility analyses) were reviewed, and academic underpinnings of NEBs were refined. Utilities adopted NEBs into marketing and targeting applications, and familiarity with NEBS grew in the field.

- Stage 4 (2008-present): In this stage, with the NEB literature and NEB estimates considerably expanded (using methods tested by many researchers), researchers revisited the issue of incorporating NEBs into regulatory benefit cost tests. Stakeholders became more comfortable with NEBs and their values, and more than a dozen states began to incorporate subsets of NEBs into TRC and other tests for program and portfolio screening.

Figure 1.1 provides an array of information. The table includes:

- Names of the individual and subcategories of NEBs included in this analysis;
- Methods used commonly / state of the art for computing values (model, survey, etc.), in the “Measured” column;
- Relative size of the NEB values in high / medium / low terms (column labeled “Size”). For some, there has not been sufficient strong estimation work to report a value (N/A);
- Consistency of the values reported in the literature, based on whether values show reasonably small variation between high and low estimates (one asterisk or wave symbol), or whether there values are consistent within program types (two asterisks) (“consistency” column);
- Whether the NEB category would be expected to vary by program or audience or not. Some values show consistent values based very closely on the energy savings (e.g., emissions, with possible variations for peak vs. off-peak programs); for others, the NEB value would be expected to vary based on the program and measures included (comfort would be expected with a weatherization program, but less so with a CFL program); and some will vary additionally with the targeted audience (low income vs. not; sectors within commercial, etc.).[9] “Ltd” denotes limited variation, or NEBs that closely follow energy savings, P indicates those that vary by program or measure types, and T indicates those varying by target audience.

This is a robust table, but all elements are included in one place to allow easier comparisons. The individual elements are discussed in the paragraphs below.

[9] *Commercial / non-residential programs also show similar patterns.*

Figure 1.1: List of NEB Categories, Perspective, and Measurement Performance [10]

(summarizes common measurement method, relative size, and variability / consistency)

Perspective, Category, and NEB	Meas.	Size	Consis	PgmVa
UTILITY PERSPECTIVE				
Payment-related				
Carrying cost on arrearages	I	L	*	P&T
Bad Debt Write-offs	I	L	*	P&T
Reduced LI subsidy pymt/discounts	I	M	*	T
Shutoffs / Reconnects	I	L	*	P&T
Notices	I	L	*	P&T
Customer calls / collections	I	L	*	P&T
Service Related				
Emergency / safety	I	M		P
Other Primary Utility				
Insurance savings	I	L/NA		Ltd
T&D savings (usually distrib)	M	L/NA		Ltd
Fewer substations / infrastructure	NA	NA		Ltd
Power quality / reliability	NA	NA		Ltd
Other Primary Utility	NA	NA		Ltd
SOCIETAL PERSPECTIVE				
Economic	M	H	*	P
Environmental / Emissions	M	H	**	Ltd
Tax effects - unempl; tax invest. credits	M	M/NA		P
H&S equipment / fires	I	NA		P
Health Care	NA	NA		P
Social welfare indicators	NA	NA		P&T
Water / Wastewater infrastructure	NA	H/NA		P
Fish / wildlife mitigation	NA	NA		Ltd
National security	NA	NA		Ltd
Other	NA	NA		
Table Key:				
Measurement method: M=model; I-incidence x value; S=Survey; NA				
Size: High / Medium / Low / NA				
Consistency:				
** low variation / consistent across programs				
* low var / consistent within pgm type; ~ fairly consistent values				
Pgm Variation: P=Vary with pgm&measures; T=vary with target audience				
Ltd=fairly constant ratio with savings (some peak vs. base var)				

Perspective, Category, and NEB	Meas.	Size	Consist	PgmVa
PARTICIPANT PERSPECTIVE				
Water and Other bills				
Water / wastewater bill savings	I	H	*	P
Other Non-energy operating costs	NA	NA		
Financial / customer service				
Shutoffs / Reconnects	I	L	*	P&T
Bill-related calls to utility	I	L	*	P&T
Collection costs, intrusions	I	L	*	P&T
Economic Dev'p / Hardship				
Economic development (low income)	NA			T
Hardship improvement / family stability (LI	NA			T
Fewer moves (LI)	S,I			P&T
Equipment Operations				
Maintenance	S	M	*	P
Lifetime extension of equipment	S	M	*	P
Equipment functionality	S	MH	*	P
Comfort, Noise, Related				
Comfort / thermal	S	MH	*	P
Noise reduction	S	M	*	P
Light quality	S	M	*	P
Health / Safety				
Health / fewer sick days work & school	S	M	*	P
IAQ / chronic illnesses	NA			P&T
Improved safety / reduced fires / insurance	S,I	M	~/*	P
Control / Education and Contributions				
Knowledge / control over bills	S	H	*	P
Contribution to the environment	S	MH	~/*	P
Satisfaction	S	LM		P
Ability to pay other bills	S	LM		P
Home Improvements				
Property value / ease of selling	S	MH	~/*	P
Aesthetics in home	S	M		P
Home durability	NA	NA		P
Special / reliability / other				
Svc. reliability/avoid interruptions	S,I	NA		Ltd; T

1.3 Status of Measurement and Usage of NEBs

Different NEBs are most suitably measured in different ways. NEBs have been measured by three key methods (and variations) in hundreds of studies conducted over the last two decades. The broad classes of measurement methods applied include:[11]

- engineering or model-based estimates, including input-output, either third-party[12] or dedicated, locally-developed models, including economic input-output and climate change models.

[10] *Definitions of most categories are fairly self-explanatory, except, perhaps, the following. Contribution to the environment is the value that participants ascribe to being able to participate and "help the environment". Satisfaction relates to their satisfaction with the improved service from new equipment. Knowledge / control over bills relates to the improved understanding of which energy-using equipment uses the most energy that they receive from some programs (with education as a component), and the consequent improvements in their ability to control energy bills. Ability to pay other bills means the program has helped free-up funds and improve their ability to pay other bills.*

[11] *From Skumatz 1998, Skumatz et.al. 2009*

[12] *For example, IMPLAN™, which is fairly commonly used to develop estimates of net job creation or net induced economic effects.*

- incremental incidence / marginal valuation estimates including factors from a combination of direct measurement and secondary literature (e.g. fewer minutes of calls from an arrearage study valued at marginal utility staff rates, fewer fires valued using insurance tables, etc.) and

- specialized comparative / ranking surveys analyzed statistically to estimate household and business-accruing NEBs (for example, results related to comfort, maintenance hassle, school days missed, etc.). The specialized surveys early-on moved beyond traditional (and volatile) estimates based on willingness to pay and willingness to accept formulations, to academically-based comparative and ranking methods[13] that have been the standard for more than 15 years. The survey work has been widely applied, expanded, and enhanced.

Figure 1.1 also notes the most common measurement approach used for the NEB category (M for model, I for incidence; S for survey, and NA / not applicable for categories that have not been frequently measured). These measurement methods, and the elapsed years, have allowed repeated testing of methods, and review of the consistency of results and valuations. When these factors are omitted from benefit-cost testing, zero values are inherently being assigned in the computations – a procedure that fails to recognize and incorporate the progress in the field. However, two decades of research now show consistent and significant results in many NEBs categories that allow more theoretically supportable formulations and reduced bias in program and portfolio decision-making. Zero is not the correct number; monetized estimates of NEBs that improve this better support effective and efficient investment of public dollars.

NEBs measure net positive and negative effects[14] beyond energy savings that have been traditionally omitted from program valuations. Best practices have emerged in NEBs measurement and attribution.[15] Measurement of the utility impacts has been fairly straightforward and has the longest history (particularly arrearages, etc.). Considerable progress has been made in the last few years in the important and high-value areas of emission and job creation impacts. Emission impacts have been measured three main ways, using simple (using system average emissions), or complex (load dispatch) methods, or using intermediate approaches (variations of peak vs. off-peak generation fuel mixes), with strong results. Work on economic impacts using third-party models has demonstrated that job creation impacts vary widely based on program type and local economic / business type mix.[16] There has also been progress on exploring methods to assess "hardship" impacts for low income programs.[17] The participant side of NEBs has seen a great deal of research, and the variety of effects have been measured using more than a dozen variations of direct, statistical, or survey (including contingent valuation, ranking and other) approaches. After more than two decades of work on NEBs, large groups of NEBs for common program types have been measured repeatedly and with fairly consistent results. The comprehensive literature review and analysis identified several remaining key gaps in NEBs work, including NEBs for kW-based programs and behavioral programs, measure-based NEBs, and estimates

---

[13] Best practices include: Estimates should only include the incremental NEB impacts associated with higher-than-standard-efficiency measures, and NTG ratios should be applied if the NEBs are to represent attributable program impacts. Best practices include providing monetized estimates of NEBs, and expressing values in consistent units (dollars per participant building per year). Skumatz and Bordner 1997, Skumatz and Dickerson 1998; options tested in Skumatz and Gardner 2006; much literature cited in Skumatz, et. al. 2009.

[14] *Negative impacts can be interpreted as quantitative measures of "barriers"*

[15] *Established in Skumatz and Bordner 1996 and Skumatz 1997.*

[16] *For example, net job creation from weatherization programs is much higher than from appliance replacement programs; and figures are higher in areas that make insulation, etc.*

[17] *See additional information on this topic in a recent AESP Brown Bag on NEBs (Skumatz and Khawaja 2010).*

for a variety of NEBs categories including strong studies of health and other effects (as depicted in Figure 1.1. However, the figure also demonstrates that there are consistent estimates for a wide variety of NEBs (depicted with asterisks in the Figure). The asterisks point out NEBs that are consistent generally across all programs (emissions, with potential variations for peak vs. baseload programs, but otherwise consistent), vs. NEBs that are consistent within program types. More detail on this issue is provided in Section 3.3 of this report.

A review of NEB results from across the US finds that the total value of NEBs varies based on program type, measures installed, and specific NEBs included. Many programs (residential, low income, and non-residential) find NEB values are near or exceed the value of energy savings. Results tend to be higher for low income programs, but multiple measure / comprehensive programs (weatherization, etc.) also see high NEBs. Ranges for these values – in dollar and percentage terms – are presented in Figure 1.2. These results were gathered from an analysis of more than 20 studies[18] of NEBs for weatherization-type programs across the nation.[19]

These types of results help program planners recognize that payback from a participant point of view does not hinge solely on energy savings,[20] and NEBs have been used in applications from marketing, targeting, program refinement (high value measures, etc.), customer benefit-cost assessment, outreach, market progress, barriers analysis, and portfolio development / screening.

Incorporation of NEBs into marketing and targeting applications was rapid; NEBs are currently most commonly used in marketing / targeting applications. NEBs are easier topics on which to market (comfort is more appealing than payback), and the NEB results have highlighted their importance in participation decision-making. NEBs also provide valuable input into program design, with negative NEBs identifying program or measure barriers and directly monetizing the additional incentives (rebates, warranty enhancements, etc.) needed to move eligible candidates to participate. Various types of NEBs analyses also identify the need for additional training or education along the supply chain or among potential participants. NEBs usage has increasingly moved beyond marketing; in some states, regulators are using NEBs in program screening and benefit-cost tests.

However, the review of state and regulatory treatment around the country indicates that, except for a few examples where "easily measured" NEBs have been included, or where regulatory test results have

---

[18] *Note that the sets of programs were not identical for the percentage and dollar calculations (not all were translatable to the same units). Studies were based on Skumatz, et. al. 2010, augmented by additional studies from New England, New York, and elsewhere (including NMR / TetraTech 2011, Skumatz et.al. 2004, Oppenheim 2012, and others.*

[19] *The table includes ranges and "typical values" for major NEBs categories. "Typical" values were defensible values selected based on a review of mean, median, and clustering of results from multiple studies. A total of more than 20 studies were reviewed, including studies in the Northeast / Mid-Atlantic region. Some studies provided only dollar values; others provided values in terms of multiples of bill savings; other provided both. There were also variations in terms of which NEB categories were and were not included in the estimation work. Where possible, we translated to consistent units. However, the studies available for the two comparisons (percentage multiplier relative to bill savings, vs. dollar adders) were not the same. For this reason, the ranges and "typical" values will not quite translate between the two treatments. Percentage adders may be the simpler treatment, allowing computation of a multiplicative adder onto existing bill savings in the B/C ratio computation, and allows simpler translations to scaling of program sizes. However, both provide valuable information.*

[20] *Efficient measures are a "bundle" of services, not just energy savings. The author notes that it is important for program planners to recognize this factor, for instance, when they set rebate amounts for program measures. A business owner (or household) will incorporate many more factors than just energy savings in their payback decision on a measure (explicitly or implicitly). Rebates that ignore these factors (NEBs) in their computation will likely not be set efficiently to achieve their behavioral and adoption goals.*

been presented with scenarios of NEBs included,[21] there has been a hesitancy to include NEBs in Benefit-Cost tests, including the TRC, etc. Researchers have explored and debated whether the proper form of NEB inclusion should be as an "adder", readily measured, all NEBs, or some hybrid.

From a conceptual point of view, exclusion is inconsistent with the philosophical underpinnings of the TRC (the test used in Maryland), and the issue is under discussion in several locales. As measurement has improved considerably, so should the confidence in including all, or at least key, NEBs in tests weighing program impacts, benefits, and costs; however, the debate is on-going.[22]

Figure 1.2: Summary of NEBs Values by Major Category (Source: Skumatz *et.al.* 2010, updated)[23]

(Dollars are added net benefit value per household per year; percentage figures should be applied to the dollar value of kWh savings)

Subtotals by major categories Weatherization Programs	Dollar NEB Values Range Low-High		Typical Value	Percentage NEB Values Range Low-High		Typical Value
UTILITY PERSPECTIVE						
Payment-related	$2.55 -	$14.50	$6.40	1% -	14.5%	4.7%
Added if Low Income subsidies avoided	$3.00 -	$25.00	$13.00	4% -	29.0%	16.4%
Service Related	$0.10 -	$8.50	$3.25	0.1% -	2.7%	0.8%
Other Primary Utility	$0.13 -	$2.60	$1.40	2.1% -	3.3%	2.4%
TOTAL UTILITY NEBs	$5.78 -	$50.60	$24.05	7.4% -	49.5%	24.4%
UTILITY NEBs MULTIPLIER	3% -	25%	12%	0.4% -	14.8%	3.3%
SOCIETAL PERSPECTIVE						
Economic	$8.00 -	$340.00	$115.00	3.0% -	237.6%	31.1%
Environmental / Emissions	$3.00 -	$180.00	$60.00	0.7% -	57.9%	7.1%
H&S equipment / fires	$0.00 -	$0.30	$0.00	0.3% -	0.3%	0.0%
Health Care	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%
Water / Wastewater infrastructure	$1.00 -	$28.00	$15.00	0.9% -	33.1%	17.0%
TOTAL SOCIETAL NEBs	$12.00 -	$548.30	$190.00	5.0% -	329.0%	55.3%
SOCIETAL NEBs MULTIPLIER	6% -	274%	95%	4.4% -	295.5%	36.5%
PARTICIPANT PERSPECTIVE						
Water and Other bills	$2.85 -	$54.00	$15.00	4.5% -	63.4%	20.0%
Financial / customer service	$0.27 -	$36.70	$3.60	8.7% -	16.4%	3.4%
Economic Dev'p / Hardship	$0.00 -	$115.00	$75.00	26.3% -	55.3%	8.0%
Equipment Operations	$26.00 -	$127.00	$82.00	17.1% -	42.7%	28.4%
Comfort, Noise, Related	$26.00 -	$105.00	$69.00	12.2% -	51.3%	26.6%
Health / Safety	$3.02 -	$100.50	$16.50	1.5% -	59.5%	12.8%
Control / Education and Contributions	$26.25 -	$177.00	$89.75	19.8% -	72.0%	26.2%
Home Improvements	$10.50 -	$77.00	$36.00	8.3% -	38.4%	18.8%
Special / reliability / other	$0.00 -	$4.05	$0.00	0.0% -	4.8%	0.0%
TOTAL PARTICIPANT NEBs	$94.89 -	$796.25	$386.85	98.5% -	403.8%	144.1%
PARTICIPANT NEBs MULTIPLIER	47% -	398%	193%	98.5% -	403.8%	144.1%
All NEBs Multipliers: Relative to Bill Savings						
Utillity	3% -	25%	12%	0% -	15%	3%
Societal	6% -	274%	95%	4% -	296%	37%
Participant	47% -	398%	193%	99% -	404%	144%
ALL Multiplers - relative to bill savings	56% -	698%	300%	103% -	714%	184%

Table Note: Studies reviewed for dollar vs. Percent groupings were not quite the same, and studies reviewed for dollar values were based on programs of different sizes and savings per customer.

[21] *For example, including 25%, 50%, 100% of NEBs.*

[22] *Particular attention has occurred in the low income side, where some NEBs are, in fact, key program goals. Several recent studies recommended inclusion of NEBs in tests, including a detailed 2001 California study (TecMarket Works, Skumatz, and Megdal 2001).*

[23] *The interpretation follows. Comfort / noise/related would be fully reflected by adding 26% to the energy savings value in the B/C test. The assumption is that these benefits decay in the same pattern as the measures that deliver the comfort (a.k.a. in relation to measure life). Alternatively, in dollar terms (and based on a somewhat different set of program documents that reported dollar-based estimates), would allow the benefits to be reflected by adding about $69 per household (per year) in net benefits attributable from the program. Those are levels reflected from a number of studies.*

1.4 Applying NEBs to Regulatory Tests

A variety of regulatory cost-effectiveness tests are used around North America to compare the cost-effectiveness of programs and optimize program investment. They are designed to take different "perspectives". States have selected different tests for a variety of reasons; a key element of the selection should relate to the state's energy policy goals. A representative listing of tests and the states that have traditionally used the tests are provided in Figure 1.1 (and further discussed in Chapter 4). The most used tests include:

- the Total Resource Cost tests (TRC) is meant to represent the utilities and their customers;
- the Societal test, a variant of the TRC meant to represent broader social views of cost-effectiveness (adding environmental costs and potentially other elements to the test)
- the Participant test is meant to represent the perspective of the participating customers;
- the Utility Cost Test (UCT), or Program Administrator Cost Test (PAC), measures costs and benefits to the utility;
- the Ratepayer Impact Measure test (RIM), measuring impacts on rates; and many other variations.

Assuming that an unbiased (or less biased) representation of costs and benefits is the objective of the test, elements of NEBs represent changes that reduce bias and better guide investment between and among programs and within portfolios. The suitable NEBs to add (replacing zero or omitted values with estimated values) for these tests are discussed in Chapter 4 (and Figure 4.1). Symmetry is important; if costs are included, the additional net benefits (NEBs) should also be included, assuming there are estimates available.

Discussions at the state and national levels have tended to center 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 (including both utility and participant costs). The theoretical consistency of the test(s) can most easily be improved by:

- including monetized estimates of the NEBs (net positive and negative) in the TRC or Societal test computations; or
- excluding all NEBs and the costs associated with achieving the NEBs (including only the "energy portion" of measure costs); or
- using the PAC or UCT test including only costs paid by the utility.

Considering Improvements

Certainly, zero values should be replaced with monetized estimates, and transparent treatment is important.[24] However, making improvements in the computed values of tests comes down to a

[24] Historically, NEBs were omitted from the (net) benefits sides of the computations in benefit-cost tests, presumably because values were not available for many of these "hard to measure" impacts. Including both benefits and costs, potentially within the confines of the "perspective") in a parallel way is the common formula. A better treatment would be to have included all factors (or excluded some explicitly for policy reasons, not missing data reasons), and explicitly identified that values for some were not (yet) available, and identified that a proxy value of "0" was assigned. Then, as numbers became available, they could be introduced, with each reducing the inherent bias in the overall equation. Note that it would also be possible to assign "weights" to various costs and benefits (including NEBs), if that was consistent with the perspective or policy goals of the tests.

question of the costs and benefits, and the associated *improvements in accuracy* of the values or components. Utilities and regulatory agencies are struggling with how to achieve that balance; what additional NEBs categories are accurately estimated within a reasonable evaluation budget? From a practical point of view, the question should be two- or three-fold:

> *Which NEB categories are most valuable, what value range arises from its (reasonable or justifiable[25] cost) measurement, and does the inclusion of the high vs. low ends of this range of values change the benefit-cost conclusion (leading to the opposite decision to include / exclude the program or measure)?*

Many NEBs have credible and consistent values and ranges already (as identified in this paper). For those for which there is uncertainty, inexpensive first-round proxies can be developed – developing a high and low range for the monetized estimate. If the inclusion of the high and low end of the ranges result in different B/C decisions, more money might be invested in the measurement to refine the proxy calculation (assuming the program investment decision is valuable), up to just shy of the value of that potentially wrong decision (Malmgren & Skumatz 2014). Further, it is clear that investing a great deal of money to refine a small value NEB by a couple percentage points is money less well spent than refining a large NEB by the same percent. Given the parallel treatment of benefits (energy or NEBs) in the formula, the "math" of benefit cost testing might even suggest that the payback from additional NEBs analysis for a program would have better return than conducting another impact evaluation on a mature or unchanged program.

Spending money on refining key NEBs values may have a greater payback than conducting another impact evaluation study on a relatively-constant program. While the impact evaluation may change the savings estimate by a few percent, a number of key NEBs categories represent potential values that are multiple times this high in relation to the bill savings, and developing high quality NEBs estimates for many categories could be funded for much less than the traditional $100,000-$250,000 impact evaluation. For some programs, it may be worthwhile to defer impact evaluations for a year and conduct a NEB analysis at least once every few years. These concepts are a type of value-based decision-making that is basic to most any economist.

1.5 NEBs in Other States

However, we recognize that most regulators like simple rules, not multi-part decisions. To deal with this cost and accuracy issue, states that are examining this issue are taking one of several tacks:

- Incorporating a simple, conservative "adder" to the benefits. Most regulators suggest they are trying to incorporate factors related to omitted environmental or emissions effects.
- Incorporating "easy to measure" NEBs to the benefits. Several states are adopting this flexible approach – with the "easy to measure" benefits varying depending on the program (e.g. water bill savings from clothes washer programs, etc.).
- Trying to measure / include all NEBs, or the leading from among several dozen NEBs, or
- A hybrid approach, using an adder plus measuring, either easy-to-measure benefits, or as many benefits as possible outside of what is included in the adder; or incorporating a base value for

[25] *Justifiable cost would be related to the "cost of a wrong decision" about the program. An expensive program might justify much higher investment in NEBs measurement if variations in the value could sway the decision about program continuation, expansion, etc.*

program-invariant NEBs, plus a program-specific adder that incorporates important NEBs for program types (weatherization vs. education, etc.), and/or customer targets (low income, etc.).

A recent comparison of the status of states around the country, in terms of their consideration of NEBs in the regulatory environment, follows *(Skumatz et. al. 2010, updated)*. The status is, of course, constantly changing.

Figure 1.2: Comparison of NEBs Treatment in Regulatory Environment, by State *(Source: Skumatz et.al, 2009, 2010, updated)*

Regulatory / Screening Application	Utilities / regions	
Program Marketing	Fairly widespread use in utilities / states across the country	MORE AGGRESSIVE ==>
Test / Pgm Screen - adder	IA (10% elec, 7.5% gas, 1999); CO (10% adder, 25% Low Inc, 2008); OR (Carbon $15/ton; 10% adder, 2008); WA (10% adder, 2008); VT (15%+15% LI); DC (10%); NY($15 adder for carbon[26]); NW (15%); for low income (LI) or <1 (CA*, ID, OR, WA*, UT, WY, NH, NY, CT)	
Test / Pgm Screen - readily measured	MA (NEBs must be "reliable & with real economic value"; utility, prop, H&S, comfort; LI; eqpt, util, all costs of complying with foreseeable environmental regulations); CA (low income); VT (maint, eqpt replacement, LI, comfort, H&S, prop, util, societal); CO (measureable with current mkt values); NH (as adder; LI); BCHydro (maint, GHG, lifetime, product loss, productivity, floorspace); DC[27] (eqpt, comfort, H&S, prop, societal); OR (esp. C&I; carbon value on societal test, PV deferred plant extension, water / sewer savings, laundry soap); CT (LI); RI (LI; quantify util, societal; H&S, eqpt, prop, comfort); NY (LI, eqpt)	
Test / Hybrid (potential adder & measured)	CO (measureable with current mkt values); OR (esp. C&I; carbon value on societal test, PV deferred plant extension, water / sewer savings, laundry soap); DC, VT.	
Test / Pgm screen - Broad	With quantification: MA, RI. MA order / decision - becoming broader - count in res & ICI / demonstratable including survey-based (not yet econ); Broad-based inclusions of all NEBs as an official screen: not yet found.	

1.6 Recommendations for Maryland

This white paper addresses weatherization-type programs, potentially including low-income and multi-family implementations.[28] The least biasing approach would be to measure the wide array of program-specific NEBs for programs. However, the data to support that computation are not currently available, and additional funding and time would be needed for this work. The funding necessary is not large. Generally, we at least recommend including the participant NEBs that are related to the program goals, and the other NEBs that are important to the system and to regulator goals (e.g. emission, etc.).

Near term:

[26] *In addition to the DPS adder, NYSERDA presented benefit /cost computations in scenarios with various percentages of included NEBs; however programs must pass without NEBs.*
[27] *Woolf 2013 used to update DC and RI; rest updated by Skumatz*
[28] *However, only a handful of studies have addressed multifamily programs. These were incorporated into our review of studies included in this paper.*

When considering costs and benefits, the most expedient method of incorporating NEBs may be two parts – a “hybrid” approach:

- Incorporate a constant dollar or percentage “adder” to represent some categories that are similar across many types of programs and are, in a sense, energy- or production- derived (e.g. emissions, infrastructure / T&D / reliability, etc.[29]). This may be in dollar per kWh or percentage terms, applied to the energy or bill savings benefits, but percentage (applied to the savings) may be the most flexible and simplest to adapt to program changes and multiple programs.
- Reduce bias further by introducing one or more percentage or dollar factors to represent other important NEBs deriving from elements specific to the program. Low income programs may add an extra factor for arrearages; weatherization programs may have higher values than some other programs due to contributions of comfort, etc. from measures that are not elements of all programs (e.g. appliance replacement programs).[30]

This might represent a “hybrid” approach to inclusion of NEBs, allowing constant adders for consistent factors across programs, and specific, or varying factors for program- or measure- or target audience-based variations. This is the approach we used in developing the adder for Maryland.

Based on our review of the results for the programs, we recommend the following for Maryland.

- **Include utility arrearage / financial impacts**: Most arrearage studies show arrearages in the range of $20-$30[31]. Larger values arise when considering programs targeting high arrearage customers or low income customers. Carrying costs for these reductions are relevant and consistent estimates have been derived. The “carrying cost” values would be on the order of 2% of bill savings (or $2.50-$4). If the utility provides low income subsidies, an adder associated with those savings may be considered, if not already incorporated elsewhere.
- **Include societal emissions impacts:** Using a fairly simple factor, the estimate of emissions benefits for Maryland might be $0.017 per kWh. This translates to a 12% adder (based on Maryland rates of about 13.7¢/kWh, and about $22/participating household in Maryland.[32] The estimate from an array of studies is 7.1% multiplier.
- **Consider societal economic impacts**: Net economic multipliers are available from the literature; however, a tailored factor for weatherization programs is available from the literature. We recommend a factor of $690,000 per million dollars in program installation dollars for programs (a multiplier of 0.69), based on a conservative estimate from other states.[33] This was the most conservative value of the three scenarios presented in the paper.
- **Include participant comfort / noise impacts**: These factors for weatherization programs are fairly consistent. We recommend a value of 10.1% for comfort alone, or 26.6% for comfort / noise / light-related benefits accruing from most weatherization programs. Dollar value versions from other programs are $69.

---

[29] *Although we believe some of these factors belong in the base part of the hybrid adder, we have not seen a multitude of strong studies on this topic (reliability); some agencies have developed T&D line loss adders, which might be an appropriate constituent. However, in some utilities, T&D, line loss, and environmental compliance values (or some subsets) are already included in avoided cost figures for energy, and should not be double-counted in those cases.*
[30] *This would argue for varying values of this second portion of the adder – with variations in values based on the type of program, and a consistent inclusion of the first adder for all programs.*
[31] *Some are as high as $60-$100.*
[32] *See Figure 5.1 for this computation.*
[33] *Gardner and Skumatz, 2009.*

- **Include health / safety impacts**: Health and safety impact to the households represent approximately 12.8% additional value beyond bill savings to households. Dollar value estimates from studies are $16.50.
- **Consider home improvement impacts**: The impacts on housing value from repairs and upgrades benefit (and are valued by) individual households, but also neighborhoods at large (the societal portion of this impact has not been well-studied). The percentage value for these impacts is about 18.8%; the dollar value suggested is $36. Excluding aesthetics (and focusing on home value), the multiplier is 10%.
- **Consider savings on other bills**: Water bills alone are very large, accounting for 20% (from about 4% to more than 50%) of energy bill savings, depending on local water rates.

Certainly the weatherization programs deliver other benefit to households, including improved knowledge in how to control their bills, which can sometimes be an objective of the programs. These impacts are on the order of 16% as a multiplier.

Figure 1.3 Summary Table of Recommended NEBs Adders for Maryland, Short Term

Category	Discussion	Value – Somewhat Conservative	Value – Very conservative
Include utility arrearage / financial impacts	Full arrears: $20 for most; $30 for low income; if carrying costs instead, $2.50-$4 (or about 2%) Consider adding low income subsidy avoidance at 16% if appropriate	2% if carrying charges; larger if full arrears; $2.50-$4; Add 16% / $13 if low income subsidies	2% / $2.50-$4 for carrying costs ($20-30 for full arrears) (higher for low income applications);
Include societal emissions impacts:	Calculations for MD 12% adder (or 1.7 ȼ /kWh, $22/MD household)[34] ; Multiplier from literature 7% / $60;	12% adder (or 1.7ȼ /kWh, $22/MD hh) (7% from literature)	12% adder (or 1.7ȼ /kWh, $22/MD hh) (7% from literature)
Consider societal economic impacts	Multiplier from literature 31% / $60; prefer simple calculations from economic multipliers from a weatherization study[35] $690,000 per $1 million in program[36]; or add factor multiplying 0.69 times per-household cost (conservative excludes admin cost)	Multiplier of 0.69 on program expenditures less admin.	Multiplier of 0.69 times program expenditures less admin.
Include participant comfort / noise impacts:	Values from literature: 10% for comfort / $30; 26% / $69 including noise and similar impacts	26% / $69	10% / $30
Include health / safety impacts:	Values from literature: 12.6% / $16.50	13% / $16.50	13% / $16.50
Consider home improvement impacts:	The literature value for these impacts is about 18.8% / $36. Excluding aesthetics (and focusing on home value), the multiplier is 10% / $18.	19% / $36;	10% / $18 excluding aesthetics
Consider water bill savings	Values from literature: 20% / $15; range depends on program measures and local water rates.	20% / $15	20% / $15
Total All (recommended and to consider)	*Percentage items are used by adding the percentage to the energy savings value in the B/C test.* *The value in dollar terms would be incorporated by adding $x per*	Base (Emissions): 12% adder (or 1.7ȼ /kWh or $22/hh for MD) (7% from literature) Plus Wx-specific adder 80% (or $124) plus economic multiplier	Base (Emissions): 12% adder (or 1.7ȼ/kWh or $22/hh for MD) (7% from literature) Plus Wx-specific adder: 55% ($82) plus economic

[34] *See Figure 5.1 for derivation of this MD-specific multiplier.*
[35] *Gardner and Skumatz, 2009.*
[36] *Conservative approach would be to omit administrative costs.*

Category	Discussion	Value – Somewhat Conservative	Value – Very conservative
	household (per year) in net benefits attributable from the program.	0.69 times program expenditures per household (or $60 econ from lit) Plus Low income adder: 16% ($13) if low income subsidies in place	multiplier 0.69 times program expenditures per HH (or $60 econ from lit) Plus Low income adder: 16% ($13) if low income subsidies in place
Total excluding "to be considered"		Base (Emissions): 12% adder (or 1.7 ¢/kWh for MD) (7% from literature) Plus Wx-specific adder: 41% adder (or $73)	Base (Emissions): 12% adder (or 1.7 ¢/kWh for MD) (7% from literature) Plus Wx- adder: 35% adder ($49)

None of these recommendations incorporate the full values for estimated NEBs; a conservative approach was taken for the short term, incorporating less than half or a fifth of the total of typical values from categories that are regularly estimated. Research is still in progress to try to develop values and methods for additional categories, excluded from these calculations.

Medium term recommendations:

We believe many of the NEB values can be improved and tailored in very short order (and for low budget). Each of these studies is in the tens of thousands of dollars, and the total is likely $50,000-$100,000. The medium-term plan would include:

- Conducting a survey of participant households to estimate important participant side benefits for this program.
- Conducting a Maryland-based economic multiplier study, using a third-party model.
- Conducting a somewhat refined emissions study, using newest relevant factors based on Maryland's generation mix and accepted / stakeholder-approved values for tonnage values.
- Using values from the multipliers table for other key values.

Longer term recommendations:
We recommend incorporation of NEBs into future study plans, so the programs may be screened and tested without unnecessary bias. The medium-term plan would include:

- Incorporating NEB questions into process (or impact) surveys for major programs with at least every other evaluation cycle, using state of the art measurement practices. The incremental cost of the survey is very low.
- Conducting a Maryland-based economic multiplier study, using a third-party model, adapting the multipliers and affected industries to be relevant to the program modeled. Weatherization programs will have higher multipliers than single-measure programs. The studies may only be needed periodically (every five years, perhaps).
- Conducting a somewhat refined emissions study, using newest relevant factors based on Maryland's generation mix and accepted / stakeholder-approved values for tonnage values. Once generated, this model can be updated by updating factors, dollar values for tons, and generation mix. Some states have developed much more complex local emissions models; this may be considered by Maryland. The factors may be kept constant for all programs, or different factors may be generated for programs expected to affect "baseload" use vs. those targeting peak usage. This refinement can be made using either the simpler or more complex models (in the simpler case, by adjusting the generation mix).

- Consider adding arrearage studies periodically to other program evaluations and use to update figures. They are inexpensive.
- Use values from the multipliers table for other key values, but consider periodically updating values based on the literature.

1.7 Summary

Twenty years on, the time ripe for reconsideration of benefit cost tests to better represent truer and more complete lists of benefit and costs, and support more optimal program investment. There has been incremental progress – but clear and distinct progress -- toward addressing the bias inherent in tests that exclude NEBs. In addition, we see a domino effect; as one state makes progress, another directly incorporates that progress into their next round of deliberations. In the near term, *inclusion of some* NEBs is better than *exclusion* of NEBs, and progress in addressing the bias in tests is important and shouldn't be delayed further.[37] Value-based decision-making argues for investment in analysis of some key NEBs categories, with tradeoffs (or deferrals) made in studies that do not have as large a potential impact on benefit-cost results. This may be one way to address the short-term measurement questions. Ratepayers, utilities, and most of all society, will benefit from enhanced metrics (NEBs inclusion in tests) that reduce bias in the billions of dollars that are invested in energy efficiency programs across the nation.

Most importantly, NEBs use in cost-effectiveness tests suffered a chicken-and-egg problem. Use in tests lagged because there were concerns about the quality of the values. Significant investment in estimation work to develop high quality values lagged because there was very limited funding of NEBs work, since they weren't incorporated into use in applications with real value to the utilities or regulators. As NEBs have become incorporated into benefit-cost tests – and as that usage grows – the robustness of the estimates will certainly grow, leading to better and better tests with less and less bias. The proper allocation of funds among energy efficiency programs – and between generation alternatives and energy efficiency – will only improve.

[37] *This represents a problem of the "perfect" comprehensive NEBs analysis getting in the way of reducing bias by including strong NEBs values as they become available, reducing bias (perfect as the enemy of the good, according to Voltaire).*

2. INTRODUCTION AND HISTORY OF NEBs RESEARCH

Non-Energy Benefit[38] (NEB) literature has shifted in the past twenty years from unmeasured lists of possible impacts, to the recognition by regulators and program administrators of these benefits in limited uses, to serious discussions of recommendations on how best to incorporate these benefits into cost-effectiveness screening. This new treatment of NEBs is now becoming seen as a best practice for energy efficiency programs; the measurement record shows NEBs are certainly non-zero, and their omission from benefit-cost computations introduces bias into program and portfolio decision-making.

Three primary policy questions arise:

- Are NEB values measurable and reliable?
- Which values are appropriate for inclusion in the context of program screening and decision-making?
- What is practical and defensible in the near – and longer term – for the State of Maryland?

This paper addresses these topics and more, including:

- Underpinnings of NEBs, the status of NEBs research, and state-of-the-art measurement methods.
- Patterns in NEB results / values, and analysis of NEBs with strong valuations and those with remaining volatility or uncertainty.
- Current treatment of NEBs in cost-effectiveness tests around North America.
- Recommendations on NEB approaches and values for cost-effectiveness tests in the State of Maryland.

2.1 Twenty Years of NEBs Progress

For decades, researchers have recognized that a significant portion of the value of energy efficiency programs comes not only from the energy savings, but from the programs' other impacts, their Non-Energy Benefits. The literature argues that Non-Energy Benefits add value to energy efficiency programs. Over the past 20 years, NEBs research has progressed from hypothesized lists of generalized benefits that might be attributable to programs, to tentative applications to low income programs, to full-fledged estimation work of scores of categories of NEBs for literally hundreds of programs across the nation. Most recently, attention has returned to the topic of incorporating the value of NEBs into cost-effectiveness screening and recommending those methods as best practices for evolving cost-effectiveness tests in efficiency programs.   Key steps in this progress are described below:[39]

- **Stage 1 (1994-1998): Organizing NEB categories into "perspectives" and basic measurement lessons,** clearly recognizing that the induced changes (beyond energy savings) have impacts on

[38] *Non-energy benefits / NEBs is the historical name of these effects in the literature. Researchers recognize that the impacts may actually be positive or negative (hence the more recent consideration of name variations, including non-energy impacts, non-energy effective, "multiple effects", omitted effects, and others). For simplicity, and to honor the founders of the literature, we use the term NEBs, recognizing that the values being discussed are actually "net" NEBs, the net of positive and potentially negative impacts delivered by energy efficiency programs beyond the directly-intended, traditional effects of energy and energy bill savings.*

[39] *This summary is adapted and expanded from Skumatz 2013.*

three different groups, with estimation methods that recognize each group's marginal costs or savings. The beneficiary groups[40] include:

- Participants in the program (e.g. reduced building operating costs, comfort, health, worker productivity, etc.);
- Utilities or agencies delivering the program (e.g. bill payment improvements, fewer line losses, system savings, etc.); and
- Society / societal NEBs (e.g. job creation, emissions reductions, infrastructure, etc.).

A list of commonly-accepted NEB categories is provided in Figure 2.1. This categorization recognizes that the same resulting change (e.g. fewer billing-related calls, fewer shutoffs / reconnects, etc.) actually has an effect (usually beneficial) for both the utility (some minutes, valued at the marginal staff time rates) and for the household or business (the same minutes, potentially valued at the household's marginal "leisure" rate per hour, or the business staff person's marginal pay rate). Organizing into beneficiaries clarifies the valuation method, recognizes the multiple effects of the programs, and allows clearly different treatments in different applications. Utility savings may be appropriate concerns in some applications, but is unlikely to be as effective as "participating household comfort" in marketing a program to potential customers. Improvements in utility system reliability may be a NEB for consideration in quite different contexts.

Basic tenets for "best practices" in measurement coming out of this stage of research include:

- NEBs can be positive or negative;
- For evaluation, programs should only count NEBs associated with the upgraded equipment[41] and
- Programs should only count the NEBs at the level of the net-to-gross / NTG[42] factor.

During this phase, considerable research into the alternatives, pros, cons, and basics of measuring for the particularly "hard to measures (HTM)" participant-side NEBs – using specially-designed ranking surveys – were extensively piloted and published. This was an important breakthrough, because, until this point, the array of important participant side NEBs beyond "hard" values like water bill savings, were ignored and called HTM. This meant that comfort, which was suspected of being a highly influential NEB category, and many other NEBs categories, were not estimated or included in computations. Statistical survey approaches were adapted to both residential and non-residential programs, where these benefits were completely unmodeled.[43]

---

[40] *First defined in Skumatz and Bordner, 1996, Skumatz 1997.*

[41] *In program evaluation, the portion of the attribute's effects (e.g. higher comfort) should properly only be the share associated with the upgrade from what the baseline would have been (usually assumed to be standard efficiency equipment to more efficient models), unless, without the program, no change from the current equipment in the home would have occurred.*

[42] *Noted in Skumatz, et.al., 2004, Skumatz et.al. 2009. NTG is a factor that represents the ratio of the gross (engineering-type) program savings that can be properly attributed as having been influenced by the program. The factor is used to derive the savings that are attributable to the program beyond what would have happened in the absence of the program. It is the combination of two effects. Free-riders (free ridership) are participants who take the incentive, but would have bought the energy-efficient measure even without the incentive (they decrease the gross savings value), and the spillover factor represents those participants (and non-participants) who were influenced by the program to adopt energy efficiency measures or behavior but did not receive the program's incentive (they add to the gross savings value).*

[43] *The progress during this period is demonstrated in the extensive citations in TecMarketWorks, Skumatz, and Megdal 2001.*

Figure 2.1: Common NEBs Categories – Perspective, Category, and NEB

UTILITY PERSPECTIVE	**PARTICIPANT PERSPECTIVE**
Payment-related	**Other bills**
Carrying cost on arrearages	Water / wastewater bill savings
Bad Debt Write-offs	Other Non-energy operating costs
Reduced LI subsidy pymt/discounts	**Financial / customer service**
Shutoffs / Reconnects	Shutoffs / Reconnects
Notices	Bill-related calls to utility
Customer calls / collections	Collection costs, intrusions
Service Related	**Economic Development / Hardship**
Emergency / safety	Economic development (low income)
Other Primary Utility	Hardship improvement / family stability (LI)
Insurance savings	Fewer moves (LI)
T&D savings (usually distrib)	**Equipment Operations**
Fewer substations / infrastructure	Maintenance
Power quality / reliability	Lifetime extension of equipment
Other Primary Utility	Equipment functionality
SOCIETAL PERSPECTIVE	**Comfort, Noise, Related**
Economic	Comfort / thermal
Environmental / Emissions	Noise reduction
Tax effects - unempl; tax invest. credits	Light quality
H&S equipment / fires	**Health / Safety**
Health Care	Health / fewer sick days work & school
Social welfare indicators	IAQ / chronic illnesses
Water / Wastewater infrastructure	Improved safety / reduced fires / insurance (gas)
Fish / wildlife mitigation	**Control / Education and Contributions**
National security	Knowledge / control over bills
Other	Contribution to the environment
	Satisfaction
	Ability to pay other bills
	Home Improvements
	Property value / ease of selling
	Aesthetics in home
	Home durability
	Special / reliability / other
	Svc. reliability/avoid interruptions

- **Stage 2 (1998-2001): First rounds of documented derivations / estimates in multiple NEB categories, and suggested incorporation into benefit/cost tests**: Projects covering more than three dozen NEBs, for a wide variety of low income, residential, and non-residential programs, organized into beneficiary categories were conducted for some of the more progressive utilities around the country. Pioneering papers detailing this research were published in several major US conference proceedings. These papers showed estimation work focused on three main sources for the documented estimates:
    - engineering or model-based estimates, either third-party[44] or dedicated, locally-developed models, including job creation from input-output models and emissions / climate-change modeling. Some statistical / regression models were also used to measure a few isolated direct impacts.
    - incremental incidence / marginal valuation estimates including factors from secondary literature (e.g. fewer minutes of calls valued at marginal utility staff rates, and at

[44] *For example, IMPLAN™, which is fairly commonly used to develop estimates of net job creation or net induced economic effects.*

household opportunity cost rates; fewer sick days lost from work valued at marginal wage rates; fewer fires valued using insurance tables; lower financial carrying costs from reductions in bills in arrears, etc.) and

- specialized comparative / ranking surveys analyzed statistically to estimate household and business-accruing NEBs. The survey work described above has been widely applied, expanded, and enhanced.

Figure 2.2 shows the measurement methods that are most commonly applied to the various NEBs categories.

Figure 2.2: Methods Used to Measure Common NEBs categories

Method	NEB category and Beneficiary / Perspective
Arrearage studies, directly or derived	Utility: Arrearages, bad debt, shutoffs/reconnects, notices, calls / collections Participant: calls, connections/ disconnections; notices
Incidence change times value	Utility: emergency / safety, T&D savings Participant: water bill savings Societal: tax effects
Engineering / Third party models	Societal: economic, emissions
Surveys:	Participant: moving, maintenance, equipment lifetimes, equipment function, comfort, noise, light quality, sick days, satisfaction, ability to pay bills, property value / aesthetics in home, satisfaction
Not currently estimated, or few studies, or multiple methods still being tested	Utility: substations / infrastructure, power quality / reliability Societal: Health, H&S, social welfare, infrastructure, wildlife, national security Participant: Deeper health benefits; IAQ

Publications and presentations pointed out the most volatile / uncertain NEBs categories (those needing further research), and the tiny / trivial NEBs categories that didn't warrant significant additional study. In addition, the publications pointed out a variety of useful applications of NEBs (marketing, targeting, barriers analysis) and pointed out the change in program benefit-cost ratios that would result as a consequence of recognizing the NEBs (in whole or as subsets, depending on the test and application).

The most comprehensive study during this phase of NEB development work, was a project commissioned by California involving the creation of a comprehensive NEBs model and associated revised "Low Income Public Purpose Test (LIPPT)". The project developed a detailed estimation model of NEBs for low income programs for all four California investor-owned utilities (IOUs), and associated recommendations on which NEBs should / could be appropriately incorporated into a revised benefits test, specifically for low income programs. This work included a careful documentation of one or more measurement methods that could be applied to bracket and estimate each specific NEBs category, including well-documented and specific numerical calculations for each NEBs category (TekMarket Works, Skumatz, and Megdal 2001). The computations in these tables were used, adapted, and localized for a number of studies for utilities around the US.[45], [46]

[45] *These tables and studies were used and reviewed in Amann 2006, Oppenheim 2008 / 2012, Schweitzer and Tonn 2002, and others.*
[46] *The progress during this period is demonstrated in the extensive citations in TecMarketWorks, Skumatz, and Megdal 2001.*

- **Stage 3 (2002-present): Expansion of estimation work to multiple programs, enhancement of best practices, and use in marketing and other applications:** With increasing interest by program staff / utilities / interveners, many evaluation firms across the nation began to conduct NEBs studies for utilities around the nation. The work included low income programs, programs in all sectors (residential, multi-family, commercial / industrial, schools, public buildings, etc.) and on all types of programs and measures including rebates, weatherization, appliance replacement, retro-commissioning, solar programs, real time pricing, and myriad other designs. Work on methods included specialized studies drilling down on one NEB or one measure in detail (Heshong, et. al, 2000), and work disaggregating program-wide NEBs to allow association of NEB values with individual measures (Smith-McClain et.al. 2004). Other publications compared and contrasted participant valuation methods (Skumatz 2002, Skumatz and Gardner 2006) and explored enhanced measurement approaches. Extensive work proceeded in developing emissions models, and additional work using third party input output models to estimate job creation and economic impacts were published. More than 100 presentations were given at US and international conferences, including some at academic venues; the literature became considerably more robust. The progress during this phase – including both measurement approaches, and patterns in results, is the focus of Chapter 2 of this report.

  The literature identifies that monetized NEBs range from 50% to more than 300% of the annual energy bill savings (depending on program and included / omitted NEBs categories), suggesting substantial omitted effects. [47]

- **Stage 4 (2008-present): Refocusing on the role of NEBs in regulatory and benefit-cost applications**: With the NEBs literature expanded (using methods tested by many researchers), and as NEBs values became demonstrated, it became clear that NEBs represented fairly significant omitted program benefits – and that the simplistic traditional regulatory tests that excluded NEBs were biased and may not be leading to optimal program investment. Stakeholders (program staff, utilities, regulators, interveners, and evaluators) became more comfortable with NEBs, opening the door to tentative revisiting of the issue of updating benefit cost tests to incorporate more comprehensive values. Think pieces on how to better re-craft tests appeared in the literature and at conferences. This work is underway, with some states sticking with the status quo, others revisiting "adders", and some considering significant changes. The progress during this phase – including the evolving treatment and consideration of NEBs in benefit-cost tests and regulatory proceedings, is the focus of Chapter 4 of this report. [48]

Over a total of more than 20 years of literature, NEBs has developed a robust and long history of research and development of methods, testing in applications to different programs / sectors / regions, presentation at formal conferences, and associated peer and academic review. In particular, this evolution and literature development has helped move the perception of NEBs among program staff, administrators, and regulators from general unfamiliarity and skepticism; to acknowledgement that some NEBs might be "real", measureable, and useful; to a greater appetite to consider NEBs in uses beyond "soft" applications like marketing; and movement to re-consider their use in benefit/cost applications and addressing NEBs in formal, regulatory, cost-effectiveness testing.

[47] *The progress during this period is demonstrated in the extensive citations in Skumatz et. al. 2009.*
[48] *The progress during this period is demonstrated in the extensive citations in Skumatz et. al. 2010.*

3. NEBs MEASUREMENT METHODS AND REVIEW OF ESTIMATIONS TO DATE

NEBs are an array of positive and negative effects of energy efficiency programs, beyond energy and associated bill savings. Over the last 20 years, a wide range of NEBs have been identified in studies.[49] Starting with work in the mid-1990s, the literature began to explore more consistent measurement methods, and sort these benefits into three "perspectives" based on the beneficiary of the effect–the utility or agency; society at-large, and the participant.[50]

This chapter discusses NEBs categories, the estimation methods applied to NEBs (to date) and discusses the values (and ranges) that have resulted from these estimations. We also discuss NEBs with strong valuations and those with remaining volatility or uncertainty.

Figure 3.1 presents a list of the three NEB "perspectives" or beneficiaries, and the major NEB categories within those categories.

Figure 3.1: Summary of NEBs Accruing from Three Perspectives (Skumatz 2009)

	Overall Description	Key "Drivers"	Specific Examples	Uses / Applications
Utility / Agency / Ratepayer Effects	These are incremental positive or negative impacts from initiatives that affect ratepayers and utilities and reduce revenue requirements. These effects are generally valued at utility (marginal) costs. They vary by type of participant (residential, low income, commercial), by overall energy savings and peak/non-peak timing and other factors.	• Financial burden • Debt collection efforts • Emergencies and/or insurance • T&D, power quality and reliability • Subsidies and transfers	Changes in: bad debt written off; carrying costs on balances; labor and other changes from changes in bill-and collection-related calls / activities; shut-offs / reconnects; line losses from power through lines; outage frequency / duration; many others	**Current**: Few. Some used to suggest targeting of bill-payment problem customers. **Potential**: Regulatory tests.
Societal Effects	Incremental non-energy impacts from initiatives that affect the greater society beyond those attributed directly to utility/ratepayers or participants. These effects are valued as appropriate to the benefit category. They vary significantly based on local economy, generation mix, peak/non-peak program effects, and other factors.	• Economic development/job creation multiplier effects • Environmental, including emissions • Health • Tax impacts • Water and other resource use • National security	Changes in: economic output; job creation; greenhouse gas (GHG) emissions; infrastructure savings for energy, water, waste water, etc.; fish and other environmental effects; assessment of energy vulnerability, others.	**Current**: A few utilities and agencies use deemed multipliers for GHG emissions or avoided environmental effects. At least one presents fraction of environmental and economic benefits as part of "scenarios" for B/C tests and portfolio analysis. **Potential**: TRC

[49] *See TecMarket Works, Skumatz, and Megdal, 2001 for a review of the early literature.*
[50] *These perspectives might be re-ordered from the large to the small (society, utility, participant), but order does not affect the results or discussion.*

	Overall Description	Key "Drivers"	Specific Examples	Uses / Applications
Participant / "User" Effects	Incremental non-energy effects from initiatives that affect those using the energy efficient equipment, beyond energy or bill savings. These effects are valued in terms relevant to the participant. They vary by user and by program and initiative (specific measures installed, education/outreach, weather, etc.).	• Payments and collection • Education • Building stock • Health • Equipment service/productivity (comfort, maintenance, etc.) • Other utilities / resources (water, etc.)	Change in:ability to understand / control energy usage; ability to pay; time spent on bill payment/collections issues; interruptions in service (shutoff, etc.); other bills (water, etc.); roperty value; health effects; direct/indirect changes in energy "service" and stream of associated income / utility / satisfaction (productivity, comfort, light quality/quantity, noise, maintenance, lifetime, reliability, etc.), and other ("green", etc. and other.	**Current**: Program marketing (limited), project screen (limited), scenario analysis (limited); some in modified TRCs when NEBs readily measurable. **Potential**: Portfolio development, program refinement, marketing, customer B/C, B/C tests.

3.1 Major Methods Used to Estimate NEBs

Different NEBs are most suitably measured in different ways. NEBs have been measured by three key methods (and variations) in hundreds of studies conducted over the last two decades. The measurement methods applied include:[51]

- engineering or model-based estimates, either third-party[52] or dedicated, locally-developed models, including job creation from input-output models and emissions / climate-change modeling. Some statistical / regression models were also used to measure a few isolated direct impacts.

- incremental incidence / marginal valuation estimates including factors from secondary literature (e.g. fewer minutes of calls valued at marginal utility staff rates, and at household opportunity cost rates; fewer sick days lost from work valued at marginal wage rates; fewer fires valued using insurance tables; lower financial carrying costs from reductions in bills in arrears, savings in household water bills, etc.) and

- specialized comparative / ranking surveys analyzed statistically to estimate household and business-accruing NEBs (for example, results related to comfort, maintenance hassle, school days missed, etc.). The specialized surveys early moved beyond traditional (and volatile) estimates based on willingness to pay and willingness to accept formulations, to academically-based comparative and ranking methods[53] that have been the standard for more than 15 years. The survey approach work has been widely applied, expanded, and enhanced.

These techniques, and the elapsed years, have allowed repeated testing of methods, and review of the consistency of results and valuations. When these factors are omitted from benefit-cost testing, zero values are inherently being assigned in the computations – a procedure that fails to recognize and

[51] *From Skumatz 1997, Skumatz et.al. 2009*
[52] *For example, IMPLAN™, which is fairly commonly used to develop estimates of net job creation or net induced economic effects.*
[53] *Skumatz 1997; options tested in Skumatz 2002 and Skumatz and Gardner 2006; much literature cited in Skumatz et.al. 2009.*

incorporate the progress in the field. However, two decades of research now show consistent and significant results that allow more theoretically supportable formulations and reduced bias in program and portfolio decision-making.

Figure 3.2 summarizes the most common methods used to estimate NEBs categories to date. The figure shows there are satisfactory methods used for some categories; however, gaps or weaknesses remain in some areas.

The most controversial area of estimation work has been in applying surveys. Survey-based estimates for participant NEBs are the appropriate approach for many categories; temperature and draft readings may seem more accurate, but comfort is a personal assessment, and comparative / ranking surveys are the best approach for many of these types of values. A detailed assessment of the major survey-based approaches (willingness to pay, comparative approaches, conjoint and ranking and other approaches) is found in Appendix A. To date, comparative approaches – asking participants (households or businesses) to compare the value of the NEB relative to their savings or another numeraire – has been the most used and represents the state of the art in this literature. The relative factor (percent or a translation of a verbal value) is multiplied times the numeraire's dollar value to calculate the NEB value. The approach, pioneered (Skumatz and Bordner 1996, Skumatz *et. al.* 2009) in the mid-1990s, succeeds because it:

- Asks households and businesses questions they can answer fairly readily
- The quickness of the questions allow incorporation into reasonable surveys, supporting large-samples of respondents
- Derives consistent results without large outliers
- Is relatively easily calculated
- Is based on academic literature.

Figure 3.2 summarizes the major estimation methods used for NEBs work over the last 20 years. More detail on methods and results is presented in Appendices A, C, and D.

Figure 3.2: Common Estimation Methods for NEB Categories

NEBs category	Measurement method applied
UTILITY PERSPECTIVE	
Payment-related	
Carrying cost on arrearages	Arrearage study
Bad Debt Write-offs	Arrearage study
Reduced LI subsidy pymt/discounts	Calculated based on savings & reduced usage
Shutoffs / Reconnects	Derived from arrearage study work
Notices	Derived from arrearage study work
Customer calls / collections	Derived from arrearage study work
Service Related	
Emergency / safety	Incidence times value
Other Primary Utility	
Insurance savings	Few studies; some work from insurance tables
T&D savings (usually distrib)	Can be calc from avoided cost, line loss factors, savings
Fewer substations / infrastructure	Few studies
Power quality / reliability	Few studies
Other Primary Utility	Depends
SOCIETAL PERSPECTIVE	
Economic	Third party models
Environmental / Emissions	Either generation mix & emission factors or complex models
Tax effects - unempl; tax invest. credits	Few studies; some factors available
H&S equipment / fires	Few studies
Health Care	Few studies
Social welfare indicators	Definition; few studies
Water / Wastewater infrastructure	Lack of studies
Fish / wildlife mitigation	Lack of studies
National security	Lack of studies
Other	
PARTICIPANT PERSPECTIVE	
Water and Other bills	
Water / wastewater bill savings	Savings factors and local rates
Other Non-energy operating costs	Reported bills; or survey
Financial / customer service	
Shutoffs / Reconnects	Arrearage, with participant values / value of service; or survey
Bill-related calls to utility	Arrearage, with participant values; or survey
Collection costs, intrusions	Arrearage, with participant values; or survey
Economic Dev'p / Hardship	
Economic development (low income)	Lack of studies
Hardship improvement / family stabilit	Definition; few studies; some progress
Fewer moves (LI)	Survey; could analyze billing record changes w/arrearage
Equipment Operations	
Maintenance	Survey; could use financial
Lifetime extension of equipment	Survey; could use financial
Equipment functionality	Participant survey
Comfort, Noise, Related	
Comfort / thermal	Participant survey
Noise reduction	Participant survey
Light quality	Participant survey
Health / Safety	
Health / fewer sick days work & school	Participant survey to date; few studies budget for more in-depth
IAQ	Survey; ltd. health literature; approaches could improve, limited to date;
Chronic and other illnesses	Survey; ltd. health literature; approaches could improve, limited to date;
Improved safety / reduced fires / insur	Incidence times value; or surveys
Control / Education and Contributions	
Knowledge / control over bills	Participant survey; could review arrearage or bill changes
Contribution to the environment	Participant survey to date; few studies budget for more in-depth
Satisfaction	Participant survey to date; few studies budget for more in-depth
Ability to pay other bills	Participant survey; few other practical approaches
Home Improvements	
Property value / ease of selling	Investment from program; some use surveys; best would be sales but impractical
Aesthetics in home	Participant survey to date; few studies budget for more in-depth
Home durability	Overlaps
Special / reliability / other	
Transaction cost	Seldom studied
Svc. reliability/avoid interruptions	Value of service approach

3.2 Status of Estimation of NEBs to date[54]

NEBs measure net positive and negative effects[55] beyond energy savings that have traditionally omitted from program valuations. Best practices have emerged in NEBs measurement and attribution. Estimates should only include the incremental NEB impacts associated with higher-than-standard-efficiency measures, and NTG ratios should be applied if the NEBs are to represent attributable program impacts. Best practices include providing monetized estimates of NEBs, and expressing values in consistent units (dollars per participant building per year).[56] Measurement of the utility impacts has been fairly straightforward and has the longest history (particularly arrearages, etc.). Considerable progress has been made in the last few years in the important and high-value areas of emission and job creation impacts. Emission impacts have been measured three main ways, using simple (using system average emissions), or complex (load dispatch) methods, or using intermediate approaches (variations of peak vs. off-peak generation fuel mixes), with strong results. Work on economic impacts using third-party models has demonstrated that job creation impacts vary widely based on program type and local economic / business type mix.[57] There has also been progress on exploring methods to assess "hardship" impacts for low income programs.[58] The participant side of NEBs has seen a great deal of research, and the variety of effects have been measured using more than a dozen variations of direct, statistical, or survey (including contingent valuation, ranking and other) approaches.

Values

A review of NEB results from across the US finds that the total value of NEBs varies quite a bit based on program type and specific NEBs included. Many programs (residential, low income, and non-residential) find NEB values are near or exceed the value of energy savings. Results tend to be higher for low income programs, but some residential and commercial programs also see high NEBs and the results vary with patterns by program design and measures. These types of results help program planners recognize that payback from a participant point of view does not hinge solely on energy savings,[59] and NEBs have been used in applications from marketing, targeting, program refinement (high value measures, etc.), customer benefit-cost assessment, outreach, market progress, barriers analysis, and portfolio development / screening. However, the review of state and regulatory treatment around the country indicates that, except for a few examples where "easily measured" NEBs have been included, or where regulatory test results have been presented with scenarios of NEBs included,[60] there has been a hesitancy to include NEBs in Benefit-Cost tests, including the TRC, etc. Researchers have explored and debated whether the proper form of NEB inclusion should be as an "adder", readily measured, all NEBs, or some hybrid. From a conceptual point of view, exclusion is inconsistent with the philosophical underpinnings of the TRC, and the issue is under discussion in several locales. As measurement has

---

[54] *This section is derived from Skumatz et. al. 2010.*

[55] *Negative impacts can be interpreted as quantitative measures of "barriers"*

[56] *Established in Skumatz and Bordner 1996, and Skumatz 1997*

[57] *For example, net job creation from weatherization programs is much higher than from appliance replacement programs; and figures are higher in areas that make insulation, etc.*

[58] *See additional information on this topic in an AESP Brown Bag Skumatz and Khawaja 2010.*

[59] *Efficient measures are a "bundle" of services, not just energy savings. The author notes that it is important for program planners to recognize this factor, for instance, when they set rebate amounts for program measures. A business owner (or household) will incorporate many more factors than just energy savings in their payback decision on a measure (explicitly or implicitly). Rebates that ignore these factors (NEBs) in their computation will likely not be set efficiently to achieve their behavioral and adoption goals.*

[60] *For example, including 25%, 50%, 100% of NEBs.*

improved considerably, so should the confidence around including all, or at least key, NEBs in tests weighing program impacts, benefits, and costs; however, the debate is on-going.[61]

Status of Estimation of Utility-Perspective NEBs: These are indirect costs or savings to the utility and its ratepayers. They include bill payment improvements, infrastructure savings. The vast majority of initial work on NEBs in the 1990s focused on utility perspective NEBs, particularly addressing topics related to arrearage changes from low income programs. Significant impacts were attributed to the programs (an average of about 20-25% reduction in arrearages); however, when valued for the utility at carrying charges, these arrearage effects were small for each participant. Further, when compared to the values associated with other benefit categories from the societal and participant perspective, the arrearage and debt/financial benefits from programs represented a small fraction of overall NEBs. There is a fair number of utility-perspective NEBs that are not addressed in the literature. These include:

- Line loss reductions. These may be addressed within some cost-effectiveness computations, but not universally, and the values are not clearly called out as an impact of the programs.
- Time of day/capacity impacts/avoided infrastructure. This is very important. However, it may be that the estimates associated with demand response programs may currently be considered direct impacts, rather than NEBs. There are effects associated with a wide array of programs, and these indirect benefits are valuable, however, it can be debated whether they fall into NEB or energy effect categories.
- Insurance impacts. These impacts cover the utility's costs for deductibles or for self-insurance from avoided emergency incidents that may be avoided through pro-active program retrofits and other program actions.

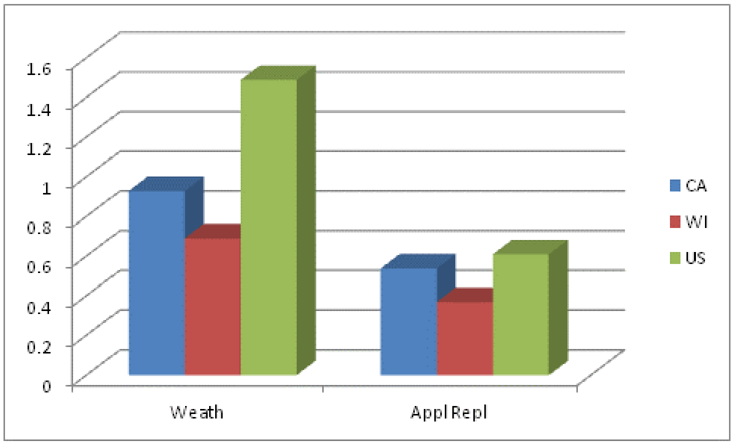
Status of Estimation of Societal-Perspective Impacts**:** These impacts are indirect program effects beyond those realized by utilities, their ratepayers, or program participants, but accrue to society at large. The literature focuses on several potential societal effects:

- **Emissions**: Consistent, defensible, and more readily-implemented modeling approaches have been developed to estimate these effects. Note that for some states (including California, moving forward), the emissions computations are addressed through avoided cost adders.
- **Job creation / economic development**: The literature shows significant impacts associated with efficiency programs which vary depending on the type of program (weatherization and education programs are more labor intensive than appliance replacement programs), region, and local industry mix. Most researchers rely on net impact analysis (from a baseline of the money spent on "generation" or on a CPI market basket using third party macroeconomic input-output models to develop these estimates, with considerable reliability.[62] An example of the impact that the territory and program type can have on the computation of estimated job impacts is presented in Figure 3.3. The study shows that labor-intensive weatherization programs create more economic multiplier effects than appliance replacement programs (for instance). Further, if the territory is expanded, the chance that the input materials are made within (and create jobs within) the territory being studied increases. Insulation may not be made in Wisconsin, but it is made in the US; on the other hand, appliances are largely not made in the US anymore.
- **Low Income Hardship benefits**: A few studies on low income programs have extended the estimation of hardship values, measuring indicators of employment scores, family stability, mobility, and reduced dependence on state benefits.

[61] *Particular attention has occurred in the low income side, where some NEBs are, in fact, key program goals. Several recent studies recommended inclusion of NEBs in tests.*
[62] *It may be argued that these"net" jobs are a cost rather than a benefit associated with the program, depending on the context.*

- **Other**: The health and safety impacts have been very sparsely studied, even though the impacts on the health care system – including incidence of chronic illnesses, etc. - may in fact be quite large. Infrastructure (water and power) and national security impacts are gaining some attention. Few other societal impacts have been seriously measured.

Figure 3.3: Net Economic Output Impact from $1 Million Program Investment (Millions of dollars) **(**Weatherization and Appliance Replacement Program examples)(Source: Gardner and Skumatz 2009)



***Status of Estimation of Participant-Perspective NEBs*:** The most controversial types of NEBs are those that accrue to the program participants. Some are measured fairly directly; water savings are computed using estimated gallon savings factors times local rates. This factor has been measured fairly regularly, but is susceptible to variation based on local water and wastewater rates. Information on financial and payments impacts (bill-related calls, shutoffs / reconnects, collection intrusions) have become fairly routine, based on the results of arrearage studies (and accompanying customer service changes) valued at household values. The status of work on the impact of effects on moves has tended to rely on surveys, although, presumably, an analysis of turnover in customer account numbers might provide indications of this value.[63] The value of reliability and time shutoff were explored early on using value of service studies (Skumatz 1998), but no new work in this area was conducted until recently. The underpinnings are based on value of service studies and reliability of T&D systems.[64] This is important on the residential side, but may be especially valuable for commercial / industrial programs. However, few NEB studies have published calculations of values.[65] Detailed analyses of household values of safety and reduced fires were conducted using insurance valuations and other approaches (Skumatz 1998, TecMarketWorks, Skumatz, and Megdal 2001) but more recent work has tended to use survey approaches for these values. Theoretically, impacts of housing value changes would be based on economic studies; these types of data have not been available.

The major source for many of the participant-side benefits has been participant comparative / ranking surveys. This has been the source for estimates and valuations of factors like operations and maintenance, comfort, productivity (for commercial / industrial), better understanding of energy use / control over bills, "doing good for the environment" and others. Some NEB lists include more than a

[63] This factor can be important, because studies (Colton 2003) note that high energy bills are an important factor leading to moving homes, and work also shows that students that change schools frequently have higher drop-out rates, affecting lifetime earnings.
[64] See Sullivan 2012, Centorella and McGranaghan 2013, EPA 2009 for valuations
[65] These types of values can also be applied to the utility reliability benefits category.

score individual benefit categories. Evaluators have tested more than eight main methods of measuring these NEBs, with the literature focusing on a relatively small subset. Each method has pros and cons, and a few studies have compared performance of different measurement methods. However, many sets of estimation results show participant NEBs often exceed the value of the energy savings from the program measures; given their: 1) significant value, 2) influence in program participation decision-making, and 3) potential role as benefits in B/C analysis. These NEBs merit continued analysis. A detailed analysis of more than a dozen methods for estimating Participant-perspective NEBs is included as Appendix A.

The analysis of dozens of NEB studies based on these types of survey sources shows fairly consistent – and generally reasonable – values for factors like comfort, noise reduction, and the other major benefit categories in residential and commercial programs. The monetized results tend to be consistent with interview-based qualitative responses on the topic of omitted benefits. The survey approaches have also been analyzed, with the literature coming to favor "comparative" approaches over other question types. The main reasons are: Respondents can answer questions phrased as comparisons better than they can answer traditional open-ended or bounded willingness-to-pay (WTP) or willingness-to-accept (WTA)-style questions; and the answers also appear to show lower variations and fewer "outlier" values.

Appendix D provides additional information on "best practices" for estimating individual categories of NEBs.

Gaps:
The comprehensive literature review and analysis identified several key gaps in NEB work. The traditional NEB analyses have neglected to include work on the association of NEBs with kilowatt-saving programs (not just kilowatt-hours). The biggest gap may be measure-based NEBs. When programs include multiple measures, few studies have examined the contribution of individual measures to the NEB values,[66] making it harder to inform programs about measures with high associated NEBs, and complicating the computation of utility benefit-costs for program screening and other applications. However, there are scores of NEB studies that specifically estimate NEBs for common types of programs. There have been NEB analyses associated with some "behavioral" and education / labeling programs (ENERGY STAR™ and a few others), but additional work is needed, because behavioral programs may likely be NEB-rich, as auxiliary benefits may be part of the "bundles of effects" recognized from adopting a new behavior.[67]

Summary:
After more than two decades of work on NEBs, large groups of NEBs for common program types have been measured repeatedly and with fairly consistent results. Gaps remain related to health impacts (societal and participant), which may emerge as high-value NEBs in the future – and are key goals of some types of EE programs (e.g. low income programs). There were also gaps in kilo-watt effects, peak, and reliability NEBs. Again, certain types of behavioral and informational programs have had virtually no NEB analysis, representing an increasingly important gap. The exceptions here are a real-time pricing analysis, and a study of commissioning.[68]

---

[66] There have been one or two exceptions that disaggregated the results by measure. See, for example, McClain-Smith, Skumatz, and Gardner 2006.
[67] Much like commissioning programs have very high NEBs, partly because the programs are designed to achieve more than just energy savings – they value and target better operation and maintenance, working together, etc.
[68] Studies by Skumatz.

3.3 Specific Value Ranges for NEBs for Weatherization Programs

Figure 3.3 summarizes an extensive review of the value ranges for a wide variety of NEBs from a review of more than 20 studies of weatherization programs.[69] The figure presents the figures in two ways:

- Dollar values – including a range from low to high from studies (excluding a few extremes that were calculated in less defensible ways), and a "typical" value, representing the average, median, or a typical value from the cluster of results.
- Percentage values – again, including high and low, but presenting the figures in terms relative to the energy or bill savings associate with the program.

There are pros and cons to presenting dollar and percentage values. Many of the dollar values are, frankly, influenced by the size, investment, or savings from the program. Some examples include financial / arrears, economics, emissions, potentially comfort, etc.). Using percentage terms for these factors may allow an easier "scale-up" when considering different programs, and also presents the estimate in terms similar to what would be used in constructing an "adder". Note that the percentage and dollar values are derived independently, and in some cases, include different numbers of studies (translations weren't possible for all studies include). Therefore, the numbers in the two sets of columns are not merely translations of each other.

Also note that these factors are estimated from weatherization programs. Certainly the types and levels of the estimated impacts are influenced by the measures included. Arrears values would very likely not be insignificant under some individual measure programs. Economics and emissions would vary, perhaps in proportion to the savings.

Appendix C provides more detailed information on the ranges for NEBs, and the estimation methods applied (Appendix A).

[69] *This includes the work by Skumatz et. al. 2010, augmented. The work covered studies from Skumatz, et. al. 2009, Oppenheim 2012, ORNL, Skumatz et. al. 2004, NMR / TetraTech 2011, and many others.*

Figure 3.4: Value Ranges from NEBs Estimates on Weatherization Programs (Source: Skumatz Economic Research Associates analysis, 2014)
(Dollars are added net benefit value per household per year; percentage figures should be applied to the dollar value of the kWh savings)

NEB Estimates from Multiple Weatherization Studies: Dollar and Percentage Analysis	Dollar NEB Values Range Low-High		Typical Value	Percent NEB Values Range Low-High		Typical Value	Notes
UTILITY PERSPECTIVE							
Payment-related							
Carrying cost on arrearages	$1.50 -	$4.00	$2.50	0.6% -	4.4%	2.0%	Total arrearages $2-$100; $20-30 typical
Bad Debt Write-offs	$0.50 -	$3.75	$1.75	0.4% -	2.0%	0.7%	
Reduced LI subsidy pymt/discounts	$3.00 -	$25.00	$13.00	3.9% -	29.0%	16.4%	IF low income program
Shutoffs / Reconnects	$0.10 -	$3.65	$0.65	0.1% -	4.4%	0.5%	
Notices	$0.05 -	$1.50	$0.60	0.1% -	1.8%	0.9%	
Customer calls / collections	$0.40 -	$1.60	$0.90	0.2% -	1.9%	0.6%	
Service Related							
Emergency / safety	$0.10 -	$8.50	$3.25	0.1% -	2.7%	0.8%	Few good studies
Other Primary Utility							
Insurance savings	$0.00 -	$0.00	$0.00	1.2% -	1.2%	1.2%	Few studies
T&D savings (usually distrib)	$0.13 -	$2.60	$1.40	0.9% -	2.1%	1.2%	Straightforward, few studies; sometimes in avoid. cost
Fewer substations / infrastructure	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Impt / needs more studies
Power quality / reliability	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Important; value of service study approach
Other Primary Utility	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	
TOTAL UTILITY NEBs	$5.78 -	$50.60	$24.05	7.4% -	49.5%	24.4%	
UTILITY NEBs MULTIPLIER	3% -	25%	12%	0.4% -	14.8%	3.3%	
SOCIETAL PERSPECTIVE							
Economic	$8.00 -	$340.00	$115.00	3.0% -	237.6%	31.1%	Better est. from expenditures & pgm type
Environmental / Emissions	$3.00 -	$180.00	$60.00	0.7% -	57.9%	7.1%	Est. based on generation mix&savings
Tax effects - unempl; tax invest. credits	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Studied early, then dropped
H&S equipment / fires	$0.00 -	$0.30	$0.00	0.3% -	0.3%	0.0%	Few studies
Health Care	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Needs studies
Social welfare indicators	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Needs studies, esp. for Low Income
Water / Wastewater infrastructure	$1.00 -	$28.00	$15.00	0.9% -	33.1%	17.0%	Needs studies
Fish / wildlife mitigation	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Little analysis to date
National security	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Little analysis to date
Other	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	
TOTAL SOCIETAL NEBs	$12.00 -	$548.30	$190.00	5.0% -	329.0%	55.3%	
SOCIETAL NEBs MULTIPLIER	6% -	274%	95%	4% -	296%	37%	

Figure 3.4: Value Ranges from NEBs Estimates on Weatherization Programs, continued (Source: Skumatz Economic Research Associates analysis, 2014)

PARTICIPANT PERSPECTIVE									
Water and Other bills									
Water / wastewater bill savings	$2.85	-	$54.00	$15.00	4.5%	-	63.4%	20.0%	Varies with regional water rates
Other Non-energy operating costs	$0.00	-	$0.00	$0.00	0.0%	-	0.0%	0.0%	
Financial / customer service									
Shutoffs / Reconnects	$0.21	-	$7.00	$1.60	0.2%	-	4.1%	1.4%	Arrears / can incorporate value of service
Bill-related calls to utility	$0.06	-	$10.00	$2.00	0.3%	-	4.0%	1.9%	
Collection costs, intrusions	$0.00	-	$19.70	$0.00	8.3%	-	8.3%	0.0%	Overlap / few separate
Economic Dev'p / Hardship						-			
Economic development (low income)	$0.00	-	$0.00	$0.00	0.0%	-	0.0%	0.0%	Needs studies
Hardship improvement / family stability (	$0.00	-	$65.00	$60.00	25.7%	-	25.9%	0.0%	Few studies
Fewer moves (LI)	$0.00	-	$50.00	$15.00	0.6%	-	29.5%	8.0%	Conservative
Equipment Operations						-			
Maintenance	$8.00	-	$43.00	$22.00	7.0%	-	9.7%	8.8%	
Lifetime extension of equipment	$7.00	-	$20.00	$20.00	3.2%	-	7.0%	5.7%	Perhaps better as financial calculation
Equipment functionality	$11.00	-	$64.00	$40.00	6.9%	-	26.0%	13.9%	
Comfort, Noise, Related									If appropriate measure types
Comfort / thermal	$12.50	-	$49.00	$30.00	3.2%	-	22.1%	10.1%	
Noise reduction	$6.75	-	$34.00	$25.00	6.0%	-	15.2%	8.5%	
Light quality	$6.75	-	$22.00	$14.00	3.0%	-	14.0%	8.0%	
Health / Safety									If appropriate measure types
Health / fewer sick days work & school	$3.00	-	$44.00	$9.00	1.4%	-	36.1%	7.4%	
IAQ	$0.00	-	$0.00	$0.00	0.0%	-	0.0%	0.0%	Few studies
Chronic and other illnesses	$0.00	-	$27.50	$0.00	0.0%	-	12.4%	0.0%	Few studies
Improved safety / reduced fires / insuran	$0.02	-	$29.00	$7.50	0.1%	-	11.0%	5.4%	

Figure 3.4: Value Ranges from NEBs Estimates on Weatherization Programs, continued (Source: Skumatz Economic Research Associates analysis, 2014)

Control / Education and Contributions									
Knowledge / control over bills	$6.75	-	$52.00	$35.00	6.0%	-	19.8%	15.7%	
Contribution to the environment	$6.00	-	$48.00	$21.75	2.8%	-	29.2%	10.6%	
Satisfaction	$13.50	-	$52.50	$33.00	0.0%	-	12.0%	0.0%	Potential overlap w/performance
Ability to pay other bills	$0.00	-	$24.50	$0.00	11.0%	-	11.0%	0.0%	Potential overlap w/bills
Home Improvements									If appropriate measure types
Property value / ease of selling	$2.50	-	$48.00	$18.00	2.3%	-	20.0%	10.0%	
Aesthetics in home	$8.00	-	$29.00	$18.00	6.0%	-	18.4%	8.8%	
Home durability	$0.00	-	$0.00	$0.00	0.0%	-	0.0%	0.0%	Possible overlap w/value
Special / reliability / other									
Transaction cost	$0.00	-	$4.05	$0.00	0.0%	-	4.8%	0.0%	Few studies/difficult concept
Svc. reliability/avoid interruptions	$0.00	-	$0.00	$0.00	0.0%	-	0.0%	0.0%	Insufficient studies; value of service; impt com'l
Other	$0.00	-	$0.00	$0.00	0.0%	-	0.0%	0.0%	
TOTAL PARTICIPANT NEBs	$94.89	-	$796.25	$386.85	98.5%	-	403.8%	144.1%	Studies estimate 89%-140% range
PARTICIPANT NEBs MULTIPLIER	47%	-	398%	193%	98.5%	-	403.8%	144.1%	excluding some terms
Subtotals - NEBs Multipliers by Type									
Relative to Participant Bill Savings									
UTILITY NEBs MULTIPLIER	3%	-	25%	12%	0%	-	15%	3%	
SOCIETAL NEBs MULTIPLIER	6%	-	274%	95%	4%	-	296%	37%	
PARTICIPANT NEBs MULTIPLIER	47%	-	398%	193%	99%	-	404%	144%	
TOTAL	56%	-	698%	300%	103%	-	714%	184%	

Table Note: Studies reviewed for dollar vs. Percent groupings were not quite the same, and studies reviewed for dollar values were based on programs of different sizes and savings per customer.

The interpretation follows. Comfort / noise/related would be fully reflected by adding 26% to the energy savings value in the B/C test. The assumption is that these benefits decay in the same pattern as the measures that deliver the comfort (a.k.a. in relation to measure life). Alternatively, in dollar terms (and based on a somewhat different set of program documents that reported dollar-based estimates), would allow the benefits to be reflected by adding about $69 per household (per year) in net benefits attributable from the program. Those are levels reflected from a number of studies.

Work beyond weatherization programs has been conducted, showing NEBs associated with individual measures or program types, including estimates for[70]:

- Refrigerators,
- Dishwashers,
- Clothes washers,
- CFL bulbs,
- CFL fixtures,
- Room air conditioners,
- Solar water heaters,
- Residential insulation,
- Boilers,
- PV,
- Commercial motors and drives,
- Commercial technical assistance,
- Commercial new construction,
- Commissioning,
- Real time pricing,
- And other programs.

A review of this research by the author confirms that participant NEBs vary in consistent and sensible patterns, with responses reflecting the services from the measures included in the program. For instance, water savings are important for clothes washers, and comfort is not a significant contributor.

3.4 Relative Uncertainties for NEB Values

In general, NEBs from various calculation results and studies have varying degrees of consistency and certainty. Some have been rarely studied; others have shown consistency from the beginnings of the estimations. Studies of patterns in NEB values across studies identified, in particular, volatility in estimates of emissions and economic results.[71] This was partly because they were large, but also because the estimation methods were not fully fleshed out.[72] Figure 3.5 addresses this issue. The most desirable cells are those on the right, in the first two rows – high value, and relatively consistent, either generally, or within certain types of programs. Some of these categories have been, or have the ability to be, well-estimated (third party models exist for emissions or economics). In addition, results for some participant survey-based categories are consistent within program types (see Figure 3.4). However, more work is especially needed in health, social indicators, etc. Some must be locally measured (water infrastructure; water rates; and possibly substation infrastructure or reliability, which may depend on the robustness of the local utility system). The asterisks in Figure 1.1 also point out NEBs that are consistent generally across all programs (emissions, with potential variations for peak vs. baseload programs, but otherwise consistent), vs. NEBs that are consistent within program types.

Figure 3.5: Variability and Patterns in NEBs Values (updated from Skumatz, *et. al.* 2010)

	Low value NEB	High value NEB
Low variation, consistent across programs		Emissions (Societal) Potentially T&D, infrastructure, reliability (utility)
Low variation WITHIN program / measure types	Utility arrearage and coll'n NEBs (utility & participant)	Economic multipliers (Societal) Home value improvement (participant; if valued according to program investment) Participant benefits including: comfort / noise / light, control over bills, equipment

[70] *Work conducted by Skumatz for various clients between 2000 and 2012*
[71] *Imbierowicz and Skumatz, ACEEE 2006.*
[72] *See Appendix D for additional discussion of methods and alternatives, particularly for emissions work. Methods for economic work have been fairly well established.*

	Low value NEB	High value NEB
		O&M / service. Safety measures, estimated using survey responses, are fairly consistent (participant)
High variation		Emergency gas service calls; emergencies; insurance (utility and participant)
Not well studied	Tax effects Wastewater / water infrastructure (unknown size) Hardship / social welfare indicators (definition; unknown size) Neighborhood property improvements (societal, unknown size) Fish / wildlife mitigation (societal, unknown size) National security (societal, unknown size)	Health and safety; health care; IAQ effects (participant and societal) Substation / infrastructure / power quality (possibly high value; utility) Reliability (participant) Fewer moves (participant)

Appendices B, C, and D provide detailed information on the ranges for NEBs and the estimations applied.

3.5 Uses of NEBs

There seems to be no shortage of informal uses or potential applications of NEBs. Introduction into more formal applications has already occurred at some level for more than a dozen states; half a dozen states have incorporated NEB elements beyond simple adders, and integration in each new state has clearly influenced adoption in the next (Malgrem and Skumatz 2014). As estimates are more and more consistent, and estimates withstand scrutiny from the range of audiences, the adoption will be expected to continue. Activity has also been undertaken to develop new protocols, encourage more transparency, and incorporate NEBs in screening procedures (including work by NEEP).

The most commonly-suggested current and potential uses of NEBs–which vary for utility, participant, and societal perspectives – are categorized in Figure 3.6 below. Enhancements on these uses are described below. The use in benefit cost tests is further discussed in the next chapter.

Figure 3.6: Summary of Current Uses for NEB Values (from Skumatz 2009)

	Utility NEBs	Societal NEBs	Participant NEBs
Marketing & targeting		Suitable	Yes
Program refinement	Yes	Yes	Yes
B/C internal customer		Suitable	Yes
Portfolio development	Yes	Yes	Yes
B/C tests	Limited; potential	Limited, high potential	Limited, High potential

NEBs provide useful information for program marketing and targeting, program refinement, and many other applications. The benefits from these qualitative and informal/informational applications have been fairly non-controversial. A discussion of the more controversial topic of how NEBs may (or may not) be adopted into program level screening and related applications is included in the next section. NEB values have been used in the following ways:

- **Program marketing / targeting**: Participant NEBs perform a function parallel to market research in product sales. NEB research uncovers those non-energy aspects of EE programs and measures that appeal to businesses and households that may be the target of the programs, and in particular to those potential participants that are not already "sold" on energy efficiency features alone. NEBs can also be used to identify high impact measures and high impact target participants for programs, optimizing impact vs. cost.

- **Program refinement**: NEBs provide feedback akin to that provided by process evaluations. Negative NEBs reflect important program barriers that can be addressed. Differences in perception of NEBs by different actors in the supply chain[73] identify information, training, or other needs at various intervention points. A detailed NEB analysis can provide information for refining the level or design of the rebate or intervention level.

- **Benefits and Costs or ROI calculations internal to the customer**: Businesses and households select equipment (and behaviors) based on an internal assessment of the benefits and costs of an array of financial and non-financial considerations and features associated with that measure or behavior. NEBs provide a mechanism for identifying and providing a financial proxy for many of these "other" features. This is a key component to understanding the participant's B/C analysis and their underlying program and participation decision-making. It provides information to refine the program and supports refinement of incentives to make the B/C ratio favorable to program objectives.[74]

- **Portfolio development**: NEB analysis allows design of portfolios that maximize societal, utility, and/or participant benefits (or targeted NEB elements) given a fixed budget. Tradeoffs can be made between programs and measures to optimize a portfolio toward an array of financial and non-financial objectives, and provides a fuller assessment of portfolio impacts.

The literature has examined the role of NEBs as important underlying motivators improving program participation, or "uptake", and demonstrated that NEB analysis along the "delivery chain" for programs can identify weak links and barriers to program implementation. In program design and evaluation, NEBs have been identified as useful in marketing and targeting; messaging; program design and refinement; incentives development, and benefit cost work. While most utilities and regulators do not treat NEBs formally, some examine them for marketing purposes. A few include "easily computed" NEBs in formal analyses (e.g., soap and water savings for washing machine programs).One utility includes percentages of NEBs in various scenarios they present to the regulators. Although NEBs have a wide array of potential applications, they have been used only sparingly by utilities and regulators around the

[73] *Termed "disconnects" (Skumatz, et. al. 2004). In research for Focus on Energy (Bensch, Skumatz and Schare 2003) the authors point out that A&E firms may be specifying and recommending fewer EE measures than owners would be willing to invest in, and that it may be leading to under-investment in EE in new construction.*

[74] *An example from a boiler program analyzed by the author illustrates this concept. Rebate levels were established to provide a customer B/C ratio that would favor the highest efficiency model. However, customers were purchasing a somewhat lower efficiency model more frequently than desired. The NEB analysis demonstrated that one of the highest value features of the other model was its small footprint, and the footprint value outweighed the difference in incentive levels. To modify behaviors, the incentives needed to be adjusted. The utility made the simplifying error of assessing customer B/C in terms of energy costs vs. purchase cost alone, rather than the greater bundle of features. NEBs provide proxies for those underlying values.*

country because of concerns about measurement uncertainty. Considerable debate has also arisen over the use–or lack of use–of NEBs in regulatory tests, and whether improved tests would lead to better program selection. NEBs may reflect some of the most important effects from energy efficiency measures and programs, and may especially represent some of the most important outcomes for low-income strategies. However, NEBs now have a history in application in many categories. The history and literature of NEBs estimation work has supported a revisiting of the use of NEBs in cost-effectiveness testing.

4. CASE FOR NEBs AND NATIONAL REVIEW OF NEBs IN COST-EFFECTIVENESS TESTING FRAMEWORK

A variety of regulatory cost-effectiveness tests are used to compare the performance of programs and optimize program investment. They have been designed to take different perspectives. States have selected different tests for a variety of reasons; a key element in the selection should reflect the State's energy policy goals. A representative listing of tests in use include:

- the Total Resource Cost tests (TRC) is meant to represent the utilities and their customers. It measures costs-effectiveness from the combined point of view of program participants and non-participants, comparing the value of avoided energy and other resources from all sources with the full cost of the efficiency measures plus all non-measure program costs;
- the Societal Test is a variant on the TRC, and is intended to represent broader social views of cost-effectiveness. The different from the TRC is the addition of environmental and other (usually a subset of) non-energy benefits and costs to the calculation.
- the Participant test is meant to represent the perspective of the participating customers. It measures cost-effectiveness from the perspective of the efficiency program participant, comparing bill savings that the customer realizes over the life of the efficiency upgrade to the cost incurred by the customer to make the upgrade (net of financial incentives the program provides);
- the Utility Cost Test (UCT), or Program Administrator Cost Test, measures costs and benefits to the utility. It compares the value of the utility's avoided costs with the cost to the utility of acquiring the efficiency resources that product those avoided costs.[75]
- the Ratepayer Impact Measure test (RIM), measures whether billing rates will go up or down as a result of an energy efficiency program, testing whether "non-participants" in a program will be better or worse off as a result of the program.

The basic formula is fairly straightforward with benefits compared to costs. However, initial formulations in protocols often omitted elements for which values weren't available; zero was substituted and/or the term left out entirely. Many of these constructs theoretically involve NEBs of varying types, based on the perspective or focus of the specific test (See Figure 4.1).

Figure 4.1: 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	Same as above plus • increased revenue	Same as above plus • Decreased revenue	AR, CO, FL, GA, HI, IA, IN, MI, MN, NC, ND, NV,	

[75] *As noted in Neme and Kushler, the primary differences from the TRC are that 1) it does not include any energy benefits for fuels that the utility does not provide; 2) it does not include any other resource benefits such as water savings; 3) it does not include any customer contributions to the cost of an efficiency investment.*

Test	Benefits	Costs	States Using Traditionally	Improved treat-ment with NEBs
Loser's Test, or non-participants test)			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 • Increases 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);[76]
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 (already generally includes all costs)
Public Purpose (PPT) (includes NEBs)	Same as above plus • Participant incentives • Quantifiable participant NEBs	Same as above	CA, KY, WI (low income)	Refined metric/ includes NEBs
Total Market Effects (TMET) (includes NEBs)	Same as above plus • Additional participant NEBs (for program and spillover participants) plus • Broader macroeconomic effects	Same as above	For evaluation purposes only	Refined metric / includes NEBs
Program Efficiency (PET) (includes NEBs)	Same as above	Same as above • Excluding participant direct costs	For evaluation purposes only	Refined metric / includes NEBs
Initial BCA (Simple BC) (includes NEBs)	Same as Public Purpose Test plus • Participant direct costs (as negative benefit) [77]	Same as above	For evaluation purposes only	Refined metric / includes NEBs

A history of omitting NEBs – admittedly because values for NEBs may have been lacking – has resulted in bias in program and portfolio decisions. The more robust nature of the literature may support incorporation going forward.

4.1 Reducing Bias in Tests: How Far to Go?

The NEBs work to date shows that the value of NEBs: 1) are not zero, and 2) can be measured. The NEBs perspectives (participant, utility, and societal) overlap all the tests. Omitting the NEBs values from the computations implicitly assumes the value is zero. This is, therefore, known to be an incorrect value, and the omission or incorrect assumption biases the benefit-cost ratio in a direction that usually leads to underinvestment or under-valuation of energy efficiency programs. Based on the results of the previous chapter, this is likely more-so for low income programs than other programs.

[76] *consistent alternative is to exclude all NEBs and costs associated with achieving them are excluded; former is easier*
[77] *Similar to the option proposed by Knight, et.al. (2006).*

Assuming that an unbiased representation of costs and benefits is the objective of the test, elements of NEBs represent changes that reduce bias and better guide investment between and among programs and within portfolios. The suitable NEBs to add (replacing the default zero value with a more robust non-zero estimate) are listed in the right hand column of Figure 4.1 above. Symmetry is important; if costs are included, the additional net benefits (NEBs) should also be included in the calculation, assuming there are estimates available.

Discussions at the state and national level have tended to center 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 (including both utility and participant costs). The theoretical consistency of the test(s) can be best improved by:

- including monetized estimates of the NEBs (net positive and negative) in the TRC or Societal test computation; or
- excluding all NEBs and the costs associated with achieving the NEBs, or
- using the UCT test including only costs paid by the utility.

Certainly, zero values should be replaced with monetized estimates, and transparent treatment of NEBs is important.[78] However, making improvements in the tests comes down to a question of the costs and benefits, and the associated *improvements in accuracy* of the values or components. Utilities and regulatory agencies are struggling with how to achieve that balance; what additional NEBs categories are accurately estimated within a reasonable evaluation budget? Frankly, from a practical point of view, the question should be two- or three-fold:

> *Which NEB categories are most valuable, what value range arises from its (reasonable or justifiable[79] cost) measurement, and does the inclusion of the high vs. low ends of this range of values change the benefit-cost conclusion (leading to the opposite decisions to include / exclude the program or measure)?*

Many NEBs have credible and consistent values and ranges (as identified in this paper). For those for which there remains uncertainty, inexpensive first-round proxies can be developed – creating a high and low range for the monetized estimate. If the inclusion of the high and low end of the ranges result in different B/C decisions, more money might be invested in the measurement to refine the calculation (assuming the program investment decision is valuable), up to just shy of the value of that potentially wrong decision (Malmgren and Skumatz 2014*)*. Further, it is clear that investing a great deal of money to refine a small value NEB by a couple percentage points is money less well spent than refining a large NEB by the same percent. Given the parallel treatment of benefits (energy or NEBs) in the formula, the "math" of benefit cost testing might even suggest that the payback from additional NEBs analysis for a

[78] Historically, NEBs were omitted from the (net) benefits sides of the computations in benefit-cost tests, presumably because values were not available for many of these "hard to measure" impacts. Including both benefits and costs, potentially within the confines of the "perspective") in a parallel way is the common formula. A better treatment would be to have included all factors (or excluded some explicitly for policy reasons, not missing data reasons) and explicitly identified that values for some were not (yet) available, and identified that a proxy value of "0" was assigned. Then, as numbers became available, they could be introduced, with each reducing the inherent bias in the overall equation. Note that it would also be possible to assign "weights" to various costs and benefits (including NEBs), if that was consistent with the perspective or policy goals of the tests.
[79] Justifiable cost would be related to the "cost of a wrong decision" about the program. An expensive program might justify much higher investment in NEBs measurement if variations in the value could sway the decision about program continuation, expansion, etc.

program would have better return than conducting another impact evaluation on a mature or unchanged program. Spending money on refining key NEBs values may have a greater payback than conducting another impact evaluation study on a relatively-constant program. While the impact evaluation may change the savings estimate by a few percent, a number of key NEBs categories represent potential values that are multiple times this high in relation to the bill savings, and developing high quality NEBs estimates for many categories could be funded for much less than the traditional $100,000-$250,000 impact evaluation. For some programs, it may be worthwhile to defer impact evaluations for a year and conduct a NEB analysis at least once every few years. These concepts are a type of value-based decision-making that is basic to most any economist.

4.2 What are Other States Doing?

However, we recognize that most regulators like simple rules, not multi-part decisions. To deal with this cost and accuracy issue, states that are examining this issue are taking one of several tacks:

- Incorporating a simple, conservative "adder" to the benefits. Most regulators suggest they are trying to incorporate factors related to omitted environmental or emissions effects.
- Incorporating "easy to measure" NEBs to the benefits. Several states are adopting this flexible approach – with the "easy to measure" benefits varying depending on the program (e.g. water bill savings from clothes washer programs, etc.).
- Trying to measure / include all NEBs, or the leading from among several dozen NEBs, or
- A hybrid approach, using an adder plus measuring either easy-to-measure benefits, or as many benefits as possible outside of what is included in the adder.

A recent comparison of the status of states around the country, in terms of their consideration of NEBs in the regulatory environment, follows (Skumatz, *et. al,* 2010, updated). The status is, of course, constantly changing.

Figure 4.2: Comparison of NEBs Treatment in Regulatory Environment, by State *(Source: Skumatz et. al., 2009, updated)*

Regulatory / Screening Application	Utilities / regions	
Program Marketing	Fairly widespread use in utilities / states across the country	MORE AGGRESSIVE ==>
Test / Pgm Screen - adder	IA (10% elec, 7.5% gas, 1999); CO (10% adder, 25% Low Inc, 2008); OR (Carbon $15/ton; 10% adder, 2008); WA (10% adder, 2008); VT (15%+15% LI); DC (10%); NY($15 adder for carbon[80]); NW (15%); for low income (LI) or <1 (CA*, ID, OR, WA*, UT, WY, NH, NY, CT)	
Test / Pgm Screen - readily measured	MA (NEBs must be "reliable & with real economic value"; utility, prop, H&S, comfort; LI; eqpt, util, all costs of complying with foreseeable environmental regulations); CA (low income); VT (maint, eqpt replacement, LI, comfort, H&S, prop, util, societal); CO (measureable with current mkt values); NH (as adder; LI); BCHydro (maint, GHG, lifetime, product loss, productivity, floorspace); DC[81] (eqpt, comfort, H&S, prop, societal); OR (esp. C&I; carbon value on societal test, PV deferred plant extension, water / sewer savings, laundry soap); CT (LI); RI (LI;	

[80] *In addition to the DPS adder, NYSERDA presented benefit /cost computations in scenarios with various percentages of included NEBs; however programs must pass without NEBs.*
[81] *Woolf 2013 used to update DC and RI; rest updated by Skumatz*

Regulatory / Screening Application	Utilities / regions	
	quantify util, societal; H&S, eqpt, prop, comfort); NY (LI, eqpt)	
Test / Hybrid (potential adder & measured)	CO (measureable with current mkt values); OR (esp. C&I; carbon value on societal test, PV deferred plant extension, water / sewer savings, laundry soap); DC, VT.	
Test / Pgm screen - Broad	With quantification: MA, RI. MA order / decision - becoming broader - count in res & ICI / demonstratable including survey-based (not yet econ); Broad-based inclusions of all NEBs as an official screen: not yet found.	

5. CONSIDERATIONS AND RECOMMENDATIONS FOR MARYLAND

5.1 Recommendations for Maryland

This white paper addresses weatherization-type programs, potentially including low-income and multi-family implementations.[82] The least biasing approach would be to measure program-specific NEBs for programs. However, the data to support that are not currently available, and additional funding and time would be needed for this work. The funding necessary is not large. Generally, we at least recommend including the participant NEBs that are related to the program goals, and the other NEBs that are important to the system and to regulator goals (e.g. emission, etc.).

Near term:
When considering costs and benefits, the most expedient method of incorporating NEBs may be two parts – a "hybrid" approach:

- Incorporate a constant dollar or percentage "adder" to represent some category – perhaps those that are similar across many types of programs and are, in a sense, energy- or production-derived (e.g. emissions[83]). This may be in dollar per kWh or percentage terms, applied to the energy or bill savings benefits, but percentage (applied to the savings) may be the most flexible and simplest to adapt to program changes and multiple programs.
- Reduce bias further by introducing one or more percentage or dollar factors to represent other important NEBs deriving from elements specific to the program. Low income programs may add an extra factor for arrearages; weatherization programs may have higher values than some other programs due to contributions of comfort, etc. from measures that are not elements of all programs (e.g. appliance replacement programs).[84]

This might represent a "hybrid" approach to inclusion of NEBs, allowing constant adders for consistent factors, and specific, or varying factors for program- or measure-based variations.

Based on our review of the results for the programs, we recommend the following for Maryland.

- **Include utility arrearage / financial impacts**: Most arrearage studies show arrearages in the range of $20-$30[85]. Larger values arise when considering programs targeting high arrearage customers or low income customers. However, a discussion is needed to determine whether the true NEB to the utility is the arrearage value itself, or (as is more commonly included in the literature) the carrying cost to the utility for these rolling arrearages. The carrying cost is only a portion (based on the interest rate), but if the assumption is made that these reduced

[82] However, only a handful of studies have addressed multifamily programs. These were incorporated into our review of studies included in this paper.
[83] Another factor that may be considered relatively invariant with respect to programs might be T&D, line loss, and similar. However, in some utilities, T&D, line loss, and environmental compliance values (or some subsets) are already included in avoided cost figures for energy, and should not be double-counted in those cases.
[84] This would argue for varying values of this second portion of the adder – with variations in values based on the type of program, and a consistent inclusion of the first adder for all programs.
[85] Some are as high as $60-$100.

arrearages also represent bad debt that the utility doesn't have to ultimately write off, or some other justification, a case might be made for using the full amount.[86] The "carrying cost" values would be on the order of 2% of bill savings (or $2.50-$4). If the utility provides low income subsidies, an adder associated with those savings may be considered, if not already incorporated elsewhere.

- **Include societal emissions impacts:** Using a fairly simple factor, the estimate of emissions benefits for Maryland might be $0.017 per kWh (which also translates to 12% adder, or about $22 per participating household per year). The estimate from an array of studies is 7.1% multiplier.
- **Consider societal economic impacts**: Net economic multipliers are available from the literature; however, the historical program values vary considerably, depending on how they are estimated, and the region assumed. The figures from Figure 5.2 or Figure 3.4 may not represent the best values, especially considering the relatively low cost of estimating these factors for weatherization programs in Maryland using third-party models. However, in the short run, we suggest considering using a factor of $690,000 per million dollars in program installation dollars for programs (a multiplier of 0.69), based on a conservative estimate from other states.[87] This is based on a study (Gardner and Skumatz 2009) that estimates a factor representing the total of net incremental labor income plus other economic output effects from an investment of $1 million in a weatherization program in the State of Wisconsin. The estimate (derived using a third-party input-output model) represents the net of the money transferring from electricity generation to the energy efficiency program. This was the most conservative of the three scenarios presented in the paper, which showed higher multipliers for California and Nationwide programs.
- **Include participant comfort / noise impacts**: These factors for weatherization programs are fairly consistent. We recommend a value of 10.1% for comfort alone, or 26.6% for comfort / noise / light-related benefits accruing from most weatherization programs. Dollar value versions from other programs are $69.
- **Include health / safety impacts**: Health and safety impact to the households represent approximately 12.8% additional value beyond bill savings to households. Dollar value estimates from studies are $16.50.
- **Consider home improvement impacts**: The impacts on housing value benefit individual households, but also neighborhoods at large (the societal portion of this impact has not been well-studied). These benefits are highly valued by the homeowners, and are one of the objectives of some low income programs (which incorporate small home repairs in some cases). The percentage value for these impacts is about 18.8%; the dollar value suggested is $36. Excluding aesthetics (and focusing on home value), the multiplier is 10%.
- **Consider savings on other bills**: Water bills alone are very large, accounting for 20% (from about 4% to more than 50%) of energy bill savings, depending on local water rates.

Applying an existing model for emissions that uses factors gleaned from the literature (SERA's "NEB-It" model), and the generation mix from Maryland, derives a set of outputs of emissions as follows. Applying dollar values provides a possible value for avoided emissions from Maryland's program.

[86] Bad debt studies are usually a different analysis.
[87] Gardner and Skumatz 2009.

The electricity generation fuel mix assumed for the analysis (from the web) is: 2.7% natural gas, 47% coal, 44.5% nuclear, 2.7% hydro, and the rest non-fossil. For simplicity, no specific accounting is made for peak vs. baseload generation mix; values would vary somewhat for peak-targeting programs.

Figure 5.1: Estimated Emissions Outputs and Values per MWh, Simplified Calculation

	GHG equivalencies, in carbon dioxide equivalencies	Pounds per MWh generated, "NEB-It" factors, avg. Maryland generation mix	Pounds per MWh generated using EIA Maryland factors
Sulfur Dioxide	0	1.805	2.3
Nitrogen Oxides	310	1.956	1.3
Carbon Dioxide	1	1054	1333
Total pounds Carbon dioxide equivalents per MWh using Maryland generation mix		1660	1736
Value per kWh saved at $X per ton CO2 10 per ton CO2 (very conservative) $20 per ton CO2 (conservative/used in remainder of report**) $100 per ton (used by environmental groups, etc.) (NOTE: Alternate values for $/ton may be selected)		0.84 cents/kWh 1.7 cents/kWh** 8.4 cents/kWh	
Multiplied times 1271 average kWh saved by MD pgm		$22/hh at $20/ton** (Alternates: CO2 ($11/hh at $10/ton CO2; $110/hh at $100/ton CO2)	
Multiplier per kWh compared to residential rates of 13.7 cents per kWh in Maryland		12% adder** (6% adder at $10/ton; 60% adder at $100/ton)	

Certainly the weatherization programs deliver other benefit to households, including improved knowledge in how to control their bills, which can sometimes be an objective of the programs. These impacts are on the order of 16% as a multiplier.

Figure 5.2 Summary Table of Recommended NEBs Adders, Short Term

Category	Discussion	Value –Somewhat Conservative	Value – Very Conservative
Include utility arrearage / financial impacts	Full arrears: $20 for most; $30 for low income; if carrying costs instead, $2.50-$4 (or about 2%) Consider adding low income subsidy avoidance at 16% if appropriate	2% if carrying charges; larger if full arrears; $2.50-$4; Add 16% / $13 if low income subsidies	2% / $2.50-$4 for carrying costs ($20-30 for full arrears) (higher for low income applications);
Include societal emissions impacts:	Multiplier from literature 7%; simple calculations 12% adder ($0.017/ kWh or $22/hh at $20/MTCO2e) for Maryland	12% ($0.017/kWh, $22/hh) MD calc; general literature: 7%; $60	12% ($0.017/kWh, $22/hh) MD calc; general literature: 7%; $60
Consider societal economic impacts	Multiplier from literature 31% / $60; prefer simple calculations from economic multipliers from a weatherization study[88] $690,000 per $1 million in program[89]; or add factor multiplying 0.69 times per-household cost.	Multiplier of 0.69 on program expenditures	Multiplier 0.69 times program expenditures less admin.
Include participant comfort / noise impacts:	Values from literature: 10% for comfort / $30; 26% / $69 including noise and similar impacts	26% / $69	10% / $30

[88] *Gardner and Skumatz, 2009.*
[89] *Conservative approach would be to omit administrative costs.*

Category	Discussion	Value –Somewhat Conservative	Value – Very Conservative
Include health / safety impacts:	Values from literature: 12.6% / $16.50	13% / $16.50	13% / $16.50
Consider home improvement impacts:	The literature value for these impacts is about 18.8% / $36. Excluding aesthetics (and focusing on home value), the multiplier is 10% / $18.	19% / $36;	10% / $18 excluding aesthetics
Consider water bill savings	Values from literature: 20% / $15; range depends on program measures and local water rates.	20% / $15	20% / $15
Total All	*Percentage items are used by adding the percentage to the energy savings value in the B/C test.* *The value in dollar terms would be incorporated by adding $x per household (per year) in net benefits attributable from the program.*	Base (Emissions): 12% adder MD (1.7¢/kWh, $22/hh MD) Plus Wx-specific adder: 80% (or $124) plus economic multiplier 0.69 times program expenditures per household (or $60) Plus Low income adder: 16% ($13) if low income subsidies in place	Base (Emissions): 12% adder MD (1.7¢/kWh, $22/hh MD) Plus Wx-specific adder: 55% ($82) plus economic multiplier 0.69 times program expenditures per HH (or $60) Plus Low income adder: 16% ($13) if low income subsidies in place
Total excluding "consider"		Base (Emissions): 12% adder for MD (or $22 MD) Plus Wx-specific adder: 41% adder (or $73)	Base (Emissions): 12% adder for MD (or $22 MD) Plus Wx-specific adder: 35% adder ($49)

None of these recommendations incorporate the full values for estimated NEBs; a conservative approach was taken for the short term, incorporating less than half or a fifth of the total of typical values from categories that are regularly estimated. Research in the field is still ongoing to develop values and methods for additional categories, excluded from these calculations.

Medium term recommendations:
We believe many of the NEB values can be improved and tailored in very short order (and for low budget). Each of these studies is in the tens of thousands of dollars, and the total is likely $50,000-$100,000. The medium-term plan would include:

- Conducting a survey of participant households to estimate important participant side benefits for this program.
- Conducting a Maryland-based economic multiplier study, using a third-party model.
- Conducting a somewhat refined emissions study, using newest relevant factors based on Maryland's generation mix and accepted / stakeholder-approved values for tonnage values.
- Using values from the multipliers table for other key values.

Longer term recommendations:
We recommend incorporation of NEBs into future study plans, so the programs may be screened and tested without unnecessary bias. The medium-term plan would include:

- Incorporating NEB questions into process (or impact) surveys for major programs with at least every other evaluation cycle, using state of the art measurement practices. The incremental cost of the survey is very low.
- Conducting a Maryland-based economic multiplier study, using a third-party model, adapting the multipliers and affected industries to be relevant to the program modeled. Weatherization

programs will have higher multipliers than single-measure programs. The studies may only be needed periodically (every five years, perhaps).

- Conducting a somewhat refined emissions study, using newest relevant factors based on Maryland's generation mix and accepted / stakeholder-approved values for tonnage values. Once generated, this model can be updated by updating factors, dollar values for tons, and generation mix. Some states have developed much more complex local emissions models; this may be considered by Maryland. The factors may be kept constant for all programs, or different factors may be generated for programs expected to affect "baseload" use vs. those targeting peak usage. This refinement can be made using either the simpler or more complex models (in the simpler case, by adjusting the generation mix).
- Consider adding arrearage studies periodically to other program evaluations and use to update figures. They are inexpensive.
- Use values from the multipliers table for other key values, but consider periodically updating values based on the literature.

5.2 Lessons and Conclusions

Twenty years on, it appears the time is becoming ripe for the reconsideration of benefit cost tests that better represent truer and more complete lists of benefit and costs, and support more optimal program investment. We see incremental progress – but clear and distinct progress -- toward addressing the bias inherent in tests that exclude NEBs. Three factors have held back progress in application of NEBs into benefit-cost tests:

- Most importantly, NEBs use in cost-effectiveness tests suffered a chicken-and-egg problem. Use in tests lagged because there were concerns about the quality of the values. Significant investment in estimation work to develop high quality values lagged because there was very limited funding of NEBs work, since they weren't incorporated into use in applications with real value to the utilities or regulators.
- There wasn't agreement on quality values for all NEBs categories, and there was concern that the estimates of some NEBs might not be accurate enough. In the near term, *inclusion of some* NEBs is better than *exclusion* of NEBs; each value helps reduce bias in tests. In addition, even if a precise point estimate isn't available, if the high and low ranges of a NEB don't change the program decision, the information is improved over using a zero value (perfect as enemy of the good issues).
- NEBs are perceived as expensive to estimate.[90]

We see a domino effect; as one state makes progress, another directly incorporates that progress into their next round of deliberations. In the near term, *inclusion of some* NEBs is better than *exclusion* of NEBs, but progress in addressing the bias in tests is important and shouldn't be delayed further. Value-based decision-making may be one way to address the short-term measurement questions. Ratepayers,

---

[90] *Considering for value-based decision-making, this may not be true. Many of the most important NEBs can be incorporated into existing process evaluations with marginal cost increases. Arrearage studies are already conducted. Comparing the "bang for the buck" for possible improvements in the overall accuracy of benefit-cost tests, another impact evaluation on a mature or little-changing program might change the benefits (savings) estimate a few percent. Deferring an impact evaluation and conducting a NEBs study would lead to benefits estimate improvements and reduction of bias many times that amount, based on the "math" of a B/C test.*

utilities, and most of all society, will only benefit from enhanced metrics (NEBs inclusion in tests) that reduce bias in the billions of dollars that are invested in energy efficiency programs across the nation.

As NEBs have become incorporated into benefit-cost tests – and as that usage grows – the robustness of the estimates will certainly grow, leading to better and better tests with less and less bias. The proper allocation of funds among energy efficiency programs – and between generation alternatives and energy efficiency – will only improve.

6.0 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.

Bensch, Ingo, Lisa A. Skumatz, Ph.D., and Stuart Schare, 2003. “Training Needs Assessment For High Performance Buildings In The Commercial Sector: Office And Education Buildings”, Wisconsin Focus on Energy, Energy Center of Wisconsin, Madison, WI.

Cadmus Group, 2011. "Memo: Non-Electric Impact (NEI) Findings for the 2011 Mass Save Home Energy Services (Mass Save) Program", prepared for Gail Azulay, NSTAR.

Cape Light Compact, 2009. "Cape Light Compact 2009 Annual Report on Energy Efficiency Activities", Barnstable, MA.

Centolella, P., and M. McGranaghan, 2013. "Understanding the Value of Uninterrupted Service", for CIGRE US National Committee 2013 Grid of the Future Symposium, Paris.

Colton, Roger D., 2003. "The Economic Development Impacts of Home Energy Assistance: The Entergy States", Fisher, Sheean & Colton, August.

Cost-Effectiveness Subcommittee of the RRM Working Group and Standardization Project Team, 2002. "LIEE Program and Measure Cost Effectiveness".

CPUC Energy Division Staff (with Dr. Ed Vine), undated. "Addressing Non-Energy Benefits in the Cost-Effectiveness Framework", San Francisco, CA.

Dalhoff, Gregory, 2007. "An Update of the Impacts of Vermont's Weatherization Assistance Program", prepared for Vermont State Office of Economic Opportunity Weatherization Assistance Program.

Daykin, Elizabeth, Jessica Aiona, and Brian Hedman, 2011(?). "Picking a Standard: Implications of Differing TRC Requirements", Proceedings of the IEPEC conference.

Dimetrosky, Scott, Lisa A. Skumatz, and Dan Violette, 2004. "Low Income Assisted Multifamily Program: Market Characterization, Assessment, and Causality Study", prepared for NYSERDA, Ithaca, NY.

Energy and Environmental Economics, Inc., 2004. "Methodology and Forecast of Long Term Avoided Costs for the Evaluation of California Energy Efficiency Programs", prepared for CPUC Energy Division, San Francisco, CA.

EPA, 2009. "Assessing the Electric System Benefits of Clean Energy".

Gardner, John, and Lisa A. Skumatz, Ph.D., 2009. Economic impacts from energy efficiency programs - Variations in multiplier effects by program type and region", Proceedings of the ECEEE conference.

Heschong, Lisa, Dr. Roger Wright, and Stacia Okura, 2000. "Daylighting and Productivity: Elementary School Studies", Proceedings of the ACEEE Summer Study on Building Conference, Asilomar, CA.

Imbierowicz, Karen, and Lisa A. Skumatz, 2004. "The Most Volatile Non-Energy Benefits (NEBs): New Research Results "Homing In" on Environmental and Economic Impacts", Proceedings of the 2008 ACEEE Summer Study on Buildings, Asilomar, CA.

Knight, Robert, Loren Lutzenhiser, and Susan Lutzenhiser. 2006. "Why Comprehensive Residential Energy Efficiency Retrofits are Undervalued", Proceedings of the ACEEE Summer Study on Building Conference, Asilomar, CA.

Malmgren, Ingrid, and Lisa A. Skumatz, Ph.D., 2014, “Lessons from the Field: Practical Applications for Incorporating Non-Energy Benefits into Cost-Effectiveness Screening”, ACEEE 2014.

National Action Plan for Energy Efficiency, 2008. "Understanding Cost-Effectiveness of Energy Efficiency Programs: Best Practices, Technical Methods, and Emerging Issuse for Policy-Makers", Energy and Environmental Economics, Inc. and Regulatory Assistance Project. <www.epa.gov/eeactionplan>.

Neme, Chris, and Marty Kushler, 2010, “Is it time to Ditch the TRC? Examining Concerns with Current Practice in Benefit-Cost Analysis”, ACEEE 2010.

NMR and TetraTech, 2011, “Massachusetts Special and Cross-Sector Studies Area, Residential and Low-Income Non-Energy Impacts (NEI) Evaluation”, August 15, 2011.

NMR, 2009. "Evaluation o the Massachusetts New Homes with Energy Star® Program", prepared for Joint Management Committee, Cambridge, MA.

Oppeheimer 2012 “Initial Comment of the Low Income Energy Affordability Network; to DPU No. 11-120, Energy Efficiency Guidelines”, memo dated 1/31/11.

Oppenheim, Jerrold, and Theo MacGregor, 2008. "How to Speak the New Language of Energy Efficiency", presentation to National Energy and Utility Affordability Conference, Denver, CO.

PA Consulting Group and Skumatz Economic Research Associates, 2005. "The Non-Energy Benefits of Wisconsin's Low Income Weatherization Program: Revised Report", prepared for Wisconsin Department of Administration, Division of Energy, Madison, WI.

Schweitzer, Martin, and Bruce Tonn, 2002. "Nonenergy Benefits from the Weatherization Assistance Program: A Summary of Findings from the Recent Literature, ORNL/CON-484.

Skumatz, Lisa A., Ph.D., 1997. “Recognizing All Program Benefits: Estimating the Non-Energy Benefits of PG&E’s Venture Partners Pilot Program (VPP)”, Proceedings of the 197 Energy Evaluation Conference, Chicago, IL.

Skumatz, Lisa A., Ph.D., 2002. "Comparing Participant Valuation Results using Three Advanced Survey Measurement Techniques: New Non-Energy Benefits Computations of Participant Value", Proceedings of the ACEEE Summer Study on Building Conference, Asilomar, CA.

Skumatz, Lisa A., Ph.D., 2006. "Methods for Measuring Non-Energy Benefits and Attributing Program Effects", Proceedings of the 2006 EEDAL Conference, London, England, June 2006.

Skumatz, Lisa A., Ph.D., 2006. “Evaluating Cost-Effectiveness, Causality, Non-Energy Benefits, and Cost-Effectiveness in Multi-Family Programs: Enhanced Techniques”, Presentation at the 2006 International Energy Efficiency in Domestic Appliances and Lighting Conference.

Skumatz, Lisa A., Ph.D., 2010. Non-Energy Benefits Analysis for Xcel Energy's Low Income Energy Efficiency Programs", prepared for Xcel Energy, Minneapolis, MN.

Skumatz, Lisa A., 2013. "NEBS - What have we learned in 20 years?", presented to IEA Capturing the Multiple Benefits of Energy Efficiency, Roundtable on Energy Provider and Consumer Benefits, October 15, Ottawa.

Skumatz, Lisa A., Ph.D., Charles Bicknell, Rose Woods, Dan Violette, and Stuart Schare, 2004. "Energy Star Labeled Homes and Home Performance with Energy Star Phase 1 Market Characterization, Assessment, and Causality MCAC Report", Prepared for NYSERDA, Ithaca, NY.

Skumatz, Lisa A., Ph.D., and Rob Bordner, 1996. “Evaluation of PG&E’s Venture Partners Pilot Project Report”, Skumatz Economic Research Associates, Seattle, WA, prepared for Pacific Gas and Electric Company, San Francisco, CA.

Skumatz, Lisa A., Ph.D. and Chris Ann Dickerson, 1998, “Extra! Extra! Non-Energy Benefits of Residential Programs Swamp Load Impacts!” Proceedings of the 1998 ACEEE conference, Asilomar, CA, August.

Skumatz, Lisa A., Ph.D., and John Gardner. 2006. "Differences in the Valuation of NEBs According to Measurement Methodology: Causes and Consequences", Proceedings of the AESP Conference. Clearwater Beach, FL.

Skumatz, Lisa A., Ph.D. and John Gardner, 2006. "Non-Energy Benefits Valuation Mechanisms: Survey and Results", Presented at Western Economics Association International, San Diego, CA.

Skumatz, Lisa A., Ph.D., and Sami Khawaja, Ph.D., 2010. "AESP webinar on NEBs in Low Income Programs", September.

Skumatz, Lisa A., Ph.D., Sami Khawaja, Ph.D. and Jane 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”, Skumatz Economic Research Associates, Prepared for California Institute for Energy and Environment (CIEE) Behavior and Energy Program, Berkeley, CA, November.

Skumatz, Lisa A., Ph.D., Sami Khawaja, Ph.D., and Richard Krop, 2010. "Non-Energy Benefits: Status, Findings, Next Steps, and Implications for Low Income Program Analyses in California", Skumatz Economic Research Associates, for SEMPRA Utilities, May.

Smith-McClain, Lisa, Lisa Skumatz, and John Gardner, 2006. "Attributing NEB Values to Specific Measures: Decomposition Results from Programs with Multiple Measures", Proceedings of the ACEEE Summer Study on Building Conference, Asilomar, CA.

Sullivan, Michael J., Ph.D., Matthew Mercurio, Josh Schellenberg, 2009. "Estimated Value of Service Reliability for Electric Utility Customers in the United States", Freeman, Sullivan & Co., prepared for Lawrence Berkeley National Laboratory, LBNL-2132E, Berkeley, CA.

Sullivan, Michael Ph.D., et. al., 2012. "Pacific Gas and Electric Company's 2012 Value of Service Study", the FSC Group, prepared for Pacific Gas and Electric Company, San Francisco, CA.

TecMarket Works, Skumatz Economic Research Associates, and Megdal and Associates. 2001. "Low Income Public Purpose Test (LIPPT) Report", Prepared for RRM Working Group Cost Effectiveness Committee, San Francisco, CA.

Wobus, Nicole et. al., 2008, "Exploring the Application of Conjoint Analysis for Estimating the Value of Non-Energy Benefits", Proceedings from the ACEEE Summer Study on Buildings, Asilomar, CA.

Woolf, Tim, Erin Malone, Jenn Kallay, Kenji Takahashi, 2013. “Energy Efficiency Cost-Effectiveness Screening in the Northeast and Mid-Atlantic States”, prepared for Regional EM&V forum, NEEP, Cambridge, MA.

See additional references in Skumatz et. al. 2009 and 2010.

APPENDIX A: SUMMARY OF PARTICIPANT NEBS MEASUREMENT METHODS

This appendix presents a summary of the major methods used to estimate participant NEBs. *See the original paper for the relevant references (Skumatz 2009).*

Figure A.1 Participant NEB Computation Approaches Proposed and Used to Date[91] (Skumatz, 2009)

Category	Description	Specific estimation approaches	Strengths	Weaknesses
A. Computational Approach / Primary Estimation:	Some categories of NEBs can be estimated fairly directly. For example, lost work time can be calculated using pre-post office records and wage rates[92] or other monetary values for time.[93] Summarily, water/sewer savings can be calculated using data on actual water and sewer rates.	1. Primary computation	• Strong, reliable, defensible results well executed	• Expensive • Lacks large sample sizes, so applicability and statistical properties are weak • Generally only used for limited number of NEB categories
B. Computation using Secondary Data Estimates:	In this case, secondary data from various sources are combined to develop a credible estimate of program impacts. For instance if secondary data are available noting risk of fires from particular measures, and the value of each average fire in terms of loss of property and life is available from, for instance, insurance companies, then these values can be multiplied times the number of measures installed to develop a total estimated value of risk from fires (or health and safety).	2. Computation from secondary sources	• Strong, reliable, defensible results • Adaptable to scenario analysis	• As strong as the secondary sources • May only be applicable to a subset of very quantitative NEB categories
C. Computation / estimation using Regression Approaches:	In some cases, statistical and regression approaches have been used to develop estimates of productivity or other effects that can be affected by confounding factors (Okura, et.al. 2000). These have been applied to several very important NEBs related to daylighting, specifically sales benefits in retail outlets, and test performance improvements in schools.	3. Regression approach	• Strong performance, with statistical reliability associated with results • Can be used with important	• Expensive, labor and skill-intensive • Data collection difficult • Can only be used to

[91] Skumatz and Gardner, "NEBs…", Western Economics Association International Paper, NV, 2004, adapted.
[92] As noted in Skumatz and Gardner, 2006, there are weaknesses from some of the direct computation methods as well. Direct computations are only available for an almost certainly non-random list of participants, and would likely be biased upward because only those businesses expecting large impacts would be likely to measure them.
[93] Some businesses may have conducted research of this type. However, estimates tend to be limited in nature, covering only the odd business or covering only one measure or a key benefit, limiting the size of the sample (and thus the error band estimation), as well as the coverage of NEBs.

Category	Description	Specific estimation approaches	Strengths	Weaknesses
			quantitative NEBs	estimate limited set of NEBs
D. Survey methods – Contingent Valuation and Willingness to Pay (WTP) / Willingness to Accept (WTA) Surveys.	Contingent valuation surveys are widely used in the environment and natural resources fields to estimate the value of intangible or hard-to-measure impacts including recreation, environmental and other effects. The contingent valuation (CV) method of non-energy benefits valuation, in its most basic form, entails simply asking respondents to estimate the value of the benefits that they experienced in dollar terms (willingness to pay WTP/ willingness to accept WTA are common approaches). An advantage of WTP surveys is that they provide specific dollar values for benefits that can be compared to each other and to the value given for the comprehensive set of program benefits. Disadvantages include the difficulty that many respondents have in answering the questions, the volatility of the responses, and significant variations in responses based on socioeconomic, demographic and attitudinal variables.[94][95]	Methods include: 4. Open-ended contingent valuation WTP / WTA questions, [96] 5. Discrete contingent valuation questions,[97] 6. Double-bounded and one-and-one-half bounded question formats,[98] 7. Ranking and ordered logit approaches[99]	• Common in literature • Clear in application • Relatively inexpensive*	• Difficult for respondents to understand and answer* • Volatile responses* • Literature cites weaknesses with open-ended responses relative to bounded options

[94] Responses to open-ended contingent valuation questions are more prone to bias (Arrow et al. 1993), and the experience of the authors has been that such responses vary more than those provided by any of the other valuation techniques discussed in this paper (Skumatz 2002, Skumatz and Gardner 2006).94 Arrow et al. (1993) list the following criticisms of the contingent valuation (CV) method for environmental valuation: 1) CV can produce results that appear to be inconsistent with assumptions of rational choice; 2) responses can seem implausibly large when considering multiple programs; 3) relatively few previous applications of the CV method have reminded respondents of relevant budget constraints; 4) it can be difficult to provide adequate background information on the programs and assume it is absorbed by respondents; 5) it can be difficult to determine "extent of market" in generating aggregate CV estimates; and 6) CV respondents may be expressing the "warm glow" of giving, rather than actual willingness to pay for the program in question

[95] Skumatz and Gardner 2006 discuss these approaches in great detail as they apply to NEBs; a summary of key issues follows. Despite the well-known limitations of direct or open-ended contingent valuation questions, there are certain situations in which they can be of use in measurement of NEBs. However, while open-ended WTP can sometimes be useful in generating a baseline, to provide more consistent and credible survey information, several variations on WTP/CV approaches can be used. 1) Discrete contingent valuation questions in which respondents are asked to give a binary "yes/no" response regarding whether they would be willing to pay a given amount for a specified good (e.g., the non-energy benefits that they experienced). This is the CV question format recommended by the 1993 NOAA panel on contingent valuation (Arrow et al. 1993). 2) Double-bounded or one-and-one-half bounded question formats, in which respondents are asked (a) to give a yes/no response to a first value, then give a follow up response to a second value, which is higher or lower depending on the response to the first question, or (b) told that the true value of the goods in question are thought to exist within a certain range, and asked to give a yes/no response to a random value, then asked to give a second response to a lower or higher value depending on the first response, unless the first response was a no to the lowest value or a yes to the highest value. These variations may increase the quality of the willingness to pay estimates obtained from referendum-type contingent valuation questions. See Cooper, Hanemann and Signorello (2002) for a discussion. 3) Ranking cards to estimate willingness to pay (also called ordered logit). The survey instrument used in this approach differs and asks respondents to rank several hypothetical scenarios in which the amount of non-energy benefits, other characteristics of the program, and a numeraire are varied at random. A rank-order logit model is then used to estimate the parameters on the utility function. The advantage to the rank-order approach is that it neither asks respondents to provide percentage or dollar estimates of the value of the non-energy benefits that they experienced nor does it ask them, hypothetically, whether predetermined values would be acceptable in exchange for those benefits. An additional advantage of this approach is that the information obtained is very robust, and the models can often be estimated with relatively small sample sizes (Weitzel and Skumatz, 2001).

Category	Description	Specific estimation approaches	Strengths	Weaknesses
	Enhancements over open-ended WTP or WTA options have been used in multiple NEB studies with varied levels of success.	[100]		
E. Survey methods – Relative scaling methods	In this approach, respondents are asked to state how much more valuable (specific or total) NEBs are relative to a base. That base may be a dollar amount, or another factor known to the respondents. Initial work focused on asking percentages higher / lower for valuations. After an extensive review of the academic literature, the use of simpler word-based comparisons (much more, etc.) could be justified and adapted, and was tested extensively.[101] The nomenclature in the academic literature for this approach is "labeled magnitude scaling" (LMS).[102]	In summary, the categories of these methods include: 8. Relative scaling in percentage terms; 9. Relative scaling in verbal terms (LMS)	• Well demonstrated in academic literature • Easy for respondents to answer / understandable* • Less volatility than WTP / WTA / CV approaches* • Inexpensive* • Can gain responses from large sample of customers, improving statistical properties	• Requires good choice of enumerative / comparison factor. • LMS requires quantitative translation from several responses
F. Ranking-	These surveys ask respondents to rank NEBs	9. AHP	• Robust	• Complex

96 Used by multiple researchers.
97 Used by multiple researchers.
98 Used in Skumatz and Gardner 2006 and other work by the authors.
99 Linked with statistical modeling approaches.
100 See Skumatz and Gardner 2004 WI and Summit Blue / Nyserda 2007.
101 The LMS was applied in Skumatz 1999. Multipliers to allow transition between words and values are presented in the literature; however, Skumatz used surveys from more than 500 respondents to confirm and refine these values for use in NEBs. The values from the academic literature were generally confirmed.
102 The relative scaling method of non-energy benefits valuation is a stated preferences approach in which survey respondents are asked to express the value of the non-energy benefits that they experienced relative to a well-understood numeraire, such as the energy savings due to the energy-efficiency measures installed through the program, program costs, or potentially any of a host of outside / non-program factors (the use of this technique and this numeraire for application to energy efficiency programs was pioneered in Skumatz and Dickerson 1997). There are several variations on the basic approach. In the direct scaling variant, respondents are asked to estimate their non-energy benefits (both positive and negative) as a percentage of their cost savings on energy. In the Labeled Magnitude Scaling (LMS) variant, respondents are asked to rate their non-energy benefits as being more valuable, less valuable or as valuable as the numeraire (e.g., their energy savings). Responses are then scaled using multipliers derived from academic sources modified by extensive empirical work from energy surveys. The relative scaling method has several advantages for use in survey research. First, program participants often find it difficult to express non-energy benefits, which are intertwined with more directly energy-related aspects of the efficiency measures that they receive, in absolute levels. However, as participants in energy efficiency programs, they are often well-attuned to changes in household or business energy costs, and therefore fully cognizant of the value of reduced energy use. Expressing the value of non-energy benefits relative to more obvious energy savings is a natural comparison that most respondents can easily make (Skumatz and Gardner 2006). As noted in Amann (2006), Skumatz pioneered this approach for NEB use and applied it in studies of residential appliance and low-income weatherization programs (Skumatz and Dickerson 1998; Skumatz, Dickerson and Coates 2000) and has since applied it in studies of ENERGY STAR home performance, new homes, and appliance programs (Fuchs, Skumatz and Ellefsen 2004). In these studies, respondents found the relative scaling questions much easier to answer than WTP questions and the responses were more consistent than those from WTP surveys.

Category	Description	Specific estimation approaches	Strengths	Weaknesses
Based Survey Approaches	or measures with alternative sets of NEBs on a two-way comparison basis (for example Analytic Hierarchy Process, AHP) or more numerous options in rank order (usually ordered logit or similar approaches). To make the estimates most robust with the least cards or questions, careful statistical design is needed (for example orthogonal models like latin squares). These approaches use information from the rankings to compute values and preferences. (Skumatz and Gardner 2004, Khawaja 2009, Wobus et.al. 2007)	10. Ranking and ordered logit approaches[103] [104]	estimates with good statistical properties are derived using this method • Requires less "monetizing" of NEBs by respondents • Strong academic grounding	question and experimental design • Can require complicated comparisons by respondents • Slower than other responses. • More difficult than some other approaches for analyzing multiple NEBs, measures.
G. Other Survey-Based Approaches - Hedonic Regression:	Most of the other methods presented have been the stated preference variety used for non-market (including environmental) goods; they require program participants to directly disclose, in one way or another, their preferences for non-energy benefits. Many non-energy benefits, however, are market goods. They are purchased by consumers, bundled with the energy-efficiency appliances that produce them, and hedonic regression approaches are suitable for these applications, decomposing price of a good as a function of its characteristics (Griliches 1961, Shelper 2001). With some variations, hedonic methods have been applied to NEBs.[105] [106]	10. Hedonic decomposition	• Well demonstrated in academic literature • Provides strong statistical and explanatory power / causal factors	• Expensive, labor and skill-intensive • Data collection complicated • Can only be used to estimate limited set of (quantitative) NEBs

[103] Linked with statistical modeling approaches.
[104] See Skumatz and Gardner 2004, Khawaja (2009) and Wobus, et.al. 2007.
[105] Because many of the characteristics of goods that give rise to non-energy benefits are abstract and subjective (e.g., light quality), the traditional hedonic regression approach may be difficult to apply. However, using the more restrictive definition of non-energy benefits, a hedonic approach to the estimation of the non-energy benefits that arise due to increased levels of energy-efficiency technology is possible and has been used. Caroll (2005) discusses a similar approach, suggesting statistical analysis of revealed preferences. Revealed preference models using a combination of program data and survey results can be used to derive estimates of NEB value. The models are used to determine how reported intent translates into action, incorporating information on, for example, the cost of the installed measures, the NEBs reported by participants, and the value of those NEBs as determined through a CV survey to derive estimates of the actual costs participants paid for the energy and NEBs associated with common projects or measures (Carroll 2005). One drawback of this approach is the time and expense associated with data collection and analysis. Skumatz and Gardner 2005 used the hedonic regressions approach to associate NEBs with specific measures in a bundled measures program.
[106] This technique may not be as robust as the stated preference approaches discussed above in that it is not capable of estimating subjective types of non-energy benefits because the more subjective characteristics of energy-using measures (aesthetics, contribution to household comfort and aesthetics, impact on health, etc.) are not available on a product-by-product basis, and are difficult to distill into readily interpretable units. This limitation notwithstanding, the hedonic regression approach non-energy benefits valuation uses data that are (a) readily available for most energy-consuming measures and (b) less susceptible to bias than direct estimates obtained from surveys. Of course, the hedonic regression approach also assumes that

Category	Description	Specific estimation approaches	Strengths	Weaknesses
H. Other survey approaches - Reported Motivations and Factor-Importance Judgments.	Customer-reported motivations for pursuing home performance projects and the relative weighting of those motivations can also be used to determine the value of the energy and non-energy benefits resulting from the project. Lutzenhiser asked customers in a California project about their motivations for buying comprehensive home performance retrofits. The reported multiple motivations among six categories (in order of frequency): specific system/building concern; environmental health and energy costs (tied); comfort; resource conservation; and other (Lutzenhiser Associates 2004).	11. Reported Motivations	• Strong performance analytically, statistically • Easy for respondents to answer • Handles quantitative and qualitative, hard and "soft" NEBs	• Expensive, labor and skill-intensive • Data collection complicated
Key: Asterisks represent results illustrated in the performance comparisons from Skumatz 2002.				

Data Collection: Studies have used a variety of methods for collecting data to support estimation of participant NEBs, including phone, mail, web, on-site interview and email approaches, as well as detailed on-site data collection using program and business records, etc. Of course, each of these data collection methods has the usual pros and cons (relative cost, speed, length / complexity tradeoffs, etc.). However, when it comes to survey-based NEBs, phone and web approaches provide additional advantages;[107] interview and on-site data collection work best for ranking and regression-based options.

Comparison of Performance of Participant NEB Approaches

Advantages and disadvantages of these various approaches have been addressed in the literature and are summarized in the Figure above. To date, only a few studies have directly compared NEB results arising from multiple measurement methods, and these findings are incorporated into the advantages and disadvantages described in the table above. These studies used two or more computational approaches to develop estimates for one program and data collection effort. Various combinations of the studies allowed comparisons between "labeled magnitude scaling" (LMS), comparative percentage, Willingness to Accept (WTA), Willingness to Pay (WTP) results, and ranking methods. The main factors used to compare the performance included:

- credible methods/demonstrated in literature;
- ease of response by respondent /comprehension of the question by respondents;[108]
- reliability of the results;[109]
- volatility of results within studies and in comparison to others;

the characteristics of a good are the only significant determinants of its price – an assumption which may or may not be reasonable depending on the goods under investigation. (Skumatz and Gardner 2006).

[107] *These include easy skip patterns (to help shorten potentially lengthy and confusing batteries of questions) and the ability to provide greater explanations if the concepts are unclear to respondents. As costs decrease, larger samples can be accommodated, supporting better statistical properties, so this is also an advantage.*

[108] *Assumed to be at least somewhat related to or reflecting reliability of individual responses – less "guessing" involved (Skumatz 2002)*

[109] *Given the types of categories of benefits being measured, "accuracy" is difficult to assess or verify. The literature that has addressed this issue tends to relate it to the next criteria, consistency of results (across similar programs, or for the same program at different times, etc.)*

- conservative /consistent results;
- cost;
- computation clarity.

Generally, the comparative research which examined quantitative and qualitative features associated with the NEB measurement methods, found that:

- WTP and WTA results (from Group D in the Table above) were weak and volatile, and confusing to respondents (and consequently had significant no response and missing observations). Respondents were slow to answer because of the confusion, and thus, data collection was relatively expensive, especially given the quality of the data in the responses. The values were generally larger (less conservative) than responses estimated using other methods (particularly Group E);
- Comparative responses (Group E) were generally consistent across programs, and very quick for respondents to answer, supporting reasonable data collection from hundreds of respondents, which improves statistical properties. The verbal comparisons (LMS) (method 9) were quicker for respondents (than Method 8), and the factors derived from the comparison of percentage vs. LMS categories were reported to be very consistent with the values reported in the academic literature.
- All methods involving WTP, WTA, and comparative valuation approaches (within Groups D and E) supported practical computation of NEBs for more than one NEB category.
- Ranking methods (Method D, number 7) provided for slower data collection than other methods, with more missing data. The questions were more difficult to construct, and only limited comparisons could be asked in the phone format, limiting the number of NEBs that could be estimated. The results were more conservative (lower) than those derived using the comparative (LMS and percentage) methods.
- The hedonic method (group G, number 10) was flexible and the results were consistent in direction and size with *a priori* theory.

To date, the LMS is a strong performer, balancing consistency, speed/efficiency/cost, and flexibility. If only one important NEB is necessary to measure, the regression-based techniques may be well-suited to the purpose. However, more work needs to be done to cross-reference and cross-check the performance and especially consistency of the results from the various methods. Only when considerable cross-checking is provided, along with demonstrated statistical properties, will confidence build for the computation of participant NEBs – especially the "softer," but still important benefits like comfort, and other NEBs. It is recommended that additional estimation work proceeds, employing multiple measures within one study to allow cross-checking and verification. Given that the literature has touted the importance of these benefits for two decades, developing credible measurement methods is important.

APPENDIX B: TABLE OF NEB VALUES

Figure B.1 presents a table of NEB values (previously presented in Chapter 3, on one page).

Figure B.1: Analysis of Ranges and Typical Values for NEBs for Weatherization Programs

NEB Estimates from Multiple Weatherization Studies: Dollar and Percentage Analysis	Dollar NEB Values Range Low-High		Typical Value	Percent NEB Values Range Low-High		Typical Value	Notes
UTILITY PERSPECTIVE							
Payment-related							
Carrying cost on arrearages	$1.50 -	$4.00	$2.50	0.6% -	4.4%	2.0%	Total arrearages $2-$100; $20-30 typical
Bad Debt Write-offs	$0.50 -	$3.75	$1.75	0.4% -	2.0%	0.7%	
Reduced LI subsidy pymt/discounts	$3.00 -	$25.00	$13.00	3.9% -	29.0%	16.4%	IF low income program
Shutoffs / Reconnects	$0.10 -	$3.65	$0.65	0.1% -	4.4%	0.5%	
Notices	$0.05 -	$1.50	$0.60	0.1% -	1.8%	0.9%	
Customer calls / collections	$0.40 -	$1.60	$0.90	0.2% -	1.9%	0.6%	
Service Related							
Emergency / safety	$0.10 -	$8.50	$3.25	0.1% -	2.7%	0.8%	Few good studies
Other Primary Utility							
Insurance savings	$0.00 -	$0.00	$0.00	1.2% -	1.2%	1.2%	Few studies
T&D savings (usually distrib)	$0.13 -	$2.60	$1.40	0.9% -	2.1%	1.2%	Straightforward, few studies; sometimes in avoid. cost
Fewer substations / infrastructure	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Impt / needs more studies
Power quality / reliability	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Important; value of service study approach
Other Primary Utility	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	
TOTAL UTILITY NEBs	$5.78 -	$50.60	$24.05	7.4% -	49.5%	24.4%	
UTILITY NEBs MULTIPLIER	3% -	25%	12%	0.4% -	14.8%	3.3%	
SOCIETAL PERSPECTIVE							
Economic	$8.00 -	$340.00	$115.00	3.0% -	237.6%	31.1%	Better est. from expenditures & pgm type
Environmental / Emissions	$3.00 -	$180.00	$60.00	0.7% -	57.9%	7.1%	Est. based on generation mix&savings
Tax effects - unempl; tax invest. credits	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Studied early, then dropped
H&S equipment / fires	$0.00 -	$0.30	$0.00	0.3% -	0.3%	0.0%	Few studies
Health Care	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Needs studies
Social welfare indicators	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Needs studies, esp. for Low Income
Water / Wastewater infrastructure	$1.00 -	$28.00	$15.00	0.9% -	33.1%	17.0%	Needs studies
Fish / wildlife mitigation	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Little analysis to date
National security	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Little analysis to date
Other	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	
TOTAL SOCIETAL NEBs	$12.00 -	$548.30	$190.00	5.0% -	329.0%	55.3%	
SOCIETAL NEBs MULTIPLIER	6% -	274%	95%	4% -	296%	37%	
PARTICIPANT PERSPECTIVE							
Water and Other bills							
Water / wastewater bill savings	$2.85 -	$54.00	$15.00	4.5% -	63.4%	20.0%	Varies with regional water rates
Other Non-energy operating costs	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	
Financial / customer service							
Shutoffs / Reconnects	$0.21 -	$7.00	$1.60	0.2% -	4.1%	1.4%	Arrears / can incorporate value of service
Bill-related calls to utility	$0.06 -	$10.00	$2.00	0.3% -	4.0%	1.9%	
Collection costs, intrusions	$0.00 -	$19.70	$0.00	8.3% -	8.3%	0.0%	Overlap / few separate
Economic Dev'p / Hardship				-			
Economic development (low income)	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Needs studies
Hardship improvement / family stability (	$0.00 -	$65.00	$60.00	25.7% -	25.9%	0.0%	Few studies
Fewer moves (LI)	$0.00 -	$50.00	$15.00	0.6% -	29.5%	8.0%	Conservative
Equipment Operations				-			
Maintenance	$8.00 -	$43.00	$22.00	7.0% -	9.7%	8.8%	
Lifetime extension of equipment	$7.00 -	$20.00	$20.00	3.2% -	7.0%	5.7%	Perhaps better as financial calculation
Equipment functionality	$11.00 -	$64.00	$40.00	6.9% -	26.0%	13.9%	
Comfort, Noise, Related							If appropriate measure types
Comfort / thermal	$12.50 -	$49.00	$30.00	3.2% -	22.1%	10.1%	
Noise reduction	$6.75 -	$34.00	$25.00	6.0% -	15.2%	8.5%	
Light quality	$6.75 -	$22.00	$14.00	3.0% -	14.0%	8.0%	
Health / Safety							If appropriate measure types
Health / fewer sick days work & school	$3.00 -	$44.00	$9.00	1.4% -	36.1%	7.4%	
IAQ	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Few studies
Chronic and other illnesses	$0.00 -	$27.50	$0.00	0.0% -	12.4%	0.0%	Few studies
Improved safety / reduced fires / insuran	$0.02 -	$29.00	$7.50	0.1% -	11.0%	5.4%	
Control / Education and Contributions							
Knowledge / control over bills	$6.75 -	$52.00	$35.00	6.0% -	19.8%	15.7%	
Contribution to the environment	$6.00 -	$48.00	$21.75	2.8% -	29.2%	10.6%	
Satisfaction	$13.50 -	$52.50	$33.00	0.0% -	12.0%	0.0%	Potential overlap w/performance
Ability to pay other bills	$0.00 -	$24.50	$0.00	11.0% -	11.0%	0.0%	Potential overlap w/bills
Home Improvements							If appropriate measure types
Property value / ease of selling	$2.50 -	$48.00	$18.00	2.3% -	20.0%	10.0%	
Aesthetics in home	$8.00 -	$29.00	$18.00	6.0% -	18.4%	8.8%	
Home durability	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Possible overlap w/value
Special / reliability / other							
Transaction cost	$0.00 -	$4.05	$0.00	0.0% -	4.8%	0.0%	Few studies/difficult concept
Svc. reliability/avoid interruptions	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	Insufficient studies; value of service; impt com'l
Other	$0.00 -	$0.00	$0.00	0.0% -	0.0%	0.0%	
TOTAL PARTICIPANT NEBs	$94.89 -	$796.25	$386.85	98.5% -	403.8%	144.1%	Studies estimate 89%-140% range
PARTICIPANT NEBs MULTIPLIER	47% -	398%	193%	98.5% -	403.8%	144.1%	excluding some terms
Subtotals - NEBs Multipliers by Type							
Relative to Participant Bill Savings							
UTILITY NEBs MULTIPLIER	3% -	25%	12%	0% -	15%	3%	
SOCIETAL NEBs MULTIPLIER	6% -	274%	95%	4% -	296%	37%	
PARTICIPANT NEBs MULTIPLIER	47% -	398%	193%	99% -	404%	144%	
TOTAL	56% -	698%	300%	103% -	714%	184%	

APPENDIX C: NEB VALUES / PATTERNS FOR LOW INCOME PROGRAMS

A detailed review[110] of the quantitative literature on low income program NEB results is summarized in the Figure below, sorted by perspective and NEB category (Figure C.1).[111] Patterns in these results are summarized in the following section. *(See the original paper for the relevant references).*

Figure C.1: Values for NEBs for Low Income Programs for Utilities around the Country (Source: Skumatz 2010**:** color groupings indicate "perspective"; LIPPT values summarize values prior to 2000; remainder updates that literature)

ID	Perspective or NEB Category	Summary of Values (per participant / yr); Implications
	UTILITY PERSPECTIVE	
	Carrying cost on arrearages	Impact values are higher for programs targeting high arrearage customers; Most standard programs in the 20-30% impact range. Dollar values clustering around $2/participant, and $32 (several in range of $60). High estimates values are reduced into this general range when translated into annual carrying cost terms.
	Bad debt written off	Impact values usually in the 20-35% range; not many studies specifically on this feature. Values $60+ for those affected, $2 when averages across all participants.
	Shutoffs	Values on order of $2 or less for many utilities; several found very high values ($100+)
	Reconnects	Net values from pennies to $50+ reconnect charge (many did not multiply times incidence)
	Notices	Few study these separately
	Customer calls / bill or emergency-related	Values on order of $0.50.
	Other bill collection cost	Few study these separately.
	Emergency gas service calls (for gas flex connector and other programs)	Based on 2 main studies – Magouirk and Blasnik. Needs more work.
	Insurance savings	Very rarely examined
0	Transmission and distribution savings (usually distribution)	Not often separately studied; embedded in utility avoided costs for some. Rules of thumb estimated percentages for some.
1	Fewer substations, etc.	Not studied to date
2	Power quality / reliability	Not studied to date
3	Reduced subsidy payments (low income)	Very directly related to the energy savings and utility's discount rate
4	Other	Tbd
	Total Perspective Utility	Lowest of the 3 perspectives. Totals range from $4-$31/HH.
5		
6	***<u>SOCIETAL PERSPECTIVE</u>***	
7	Economic development benefits – direct and indirect multipliers	Very dependent on measures and program type.

[110] This appendix is derived from Skumatz 2010.
[111] A table summarizing the specific estimation methods used in the 2000 Low Income Public Purpose Test is presented in Appendix A.

ID	Perspective or NEB Category	Summary of Values (per participant / yr); Implications
8	Tax effects - (2 possible effects: related to unemployment and income taxes from job creation / economic development; another effect possibly related to tax credits for investment in certain measures / PV / solar, etc.)	Directly related to above plus local tax schedules. Can be calculated relatively easily. Not volatile in an unpredictable way.
9	Emissions / environmental (trading values and/or health / hazard benefits)	Dependent on fuel mix, time of day (peak / off-peak) or can use more complex algorithms. Varies by utility. For California, the values are embedded in avoided cost adders.
0	Health and safety equipment	Very few studies; presumably very dependent on measures
1	Water and waste water treatment or supply plants	Rarely or never studied
2	Fish / wildlife mitigation	Never studied
3	National security	Rarely studied
4	Health care	Rarely studied
5	Reduced dependency / Improved social indicators of family stability and employment / reduced dependence on state assistance	Rarely studied, important
6	Other	
	Total Perspective Societal	Potentially valuable when economic development and emissions effects included.
7	***HOUSEHOLD PARTICIPANT PERSPECTIVE***	
8	Water / wastewater bill savings	Somewhat valuable, especially in California with high water and sewer rates. Easily computed from secondary data; depends on measures installed. $5-12/HH/yr
9	Operating costs (non-energy)	Rarely studied.
0	Equipment maintenance	Survey-based; $17-22 estimates.
1	Equipment performance (push air better, etc.)	Many studies; important, especially with comfort; extant values $14-18
2	Equipment lifetime	Few quantitative results separate from surveys.
3	Shutoffs	Survey based or based on computations of time value. Seems to indicate small values because of low incidence. Current values vary from a few cents to $12. Varies based on procedures at utility and charges.
4	Reconnects	Same as above.
5	Property value benefits / selling	Potentially very important, but also very local and program-specific (what measures, etc.). Needs more study, but likely very hard (costly) to compute because of data collection (not because it is complex). Varies from a few dollars to more than $20.
6	(Bill-related) calls to utility	Time value of data from arrearage study. Generally around $0.30; one study finds up to $8.
7	Comfort	Valuable in almost all studies; see line 31. Up to $50+ per year in one study. Commonly one of the top benefits from low income programs.
8	Aesthetics / appearance	Survey-based; should be related to line 35
	Fires / insurance damage (gas)	Rarely studied; indirect; incidence data very thin.

ID	Perspective or NEB Category	Summary of Values (per participant / yr); Implications
9		
0	Lighting / quality of light	Survey-based; depends on measures installed. One study showed $25.
1	Noise (internal / equipment)	Survey-based; depends on measures installed; extant values $15-20.
2	Noise (external)	Same as above; extant values $13-17
3	Safety	Few incidence studies – needs more work; extant values about $20.
4	Control over bill	Survey-based historically. Values ~$30.
5	Understanding / knowledge	Needs more study. Potentially important.
6	"Care" or "hardship" (low income) - and/ or see row 53 - related	Important for further exploration.
7	Indoor air quality	Not strongly recognized as separate impact in most studies.
8	Health / lost days at work or school	Important; high value for some programs, but most between $4 and $12 / HH / yr.
9	Fewer moves	The mobility value is potentially high, but incidence studies are few. One study found value of more than $60; most use more conservative numbers and derive lower estimates (under $1 because of small incidence)
0	Doing good for environment	Highly valued by participants; not clear value to programs
1	Savings in other fuels or services (as relevant)	Direct when measuring gas and electric; not many other services studied.
2	GHG and environmental effects	Measured under societal.
3	Employment and family stability, reduced dependence on state assistance	Important; see line 46
	Other	Depends.
5	NEGATIVES include: Installation hassles / mess, negative values from items above	Not usually found to be important / valuable.
	Total Perspective Participant	Majority of value for some programs

Results, Patterns, and Conclusions from Low Income Program NEB Results

A review of these findings, allows us to examine some patterns by region and program type. Figure C.2, C.3, and C.4 summarize patterns in the results for each of the three perspectives, respectively utility, societal, and participant. Note that, in almost all cases, the values are based on an analysis of program-wide NEBs – not based on measure-specific impacts.

Figure C.2: Patterns in Utility NEBs by Program Type and Region (Source: Skumatz 2010)

	Utility NEBs
General results	Small – less than 10% of total NEBs in most cases.
Variations by Program type	The effects have historically been larger for low income programs because the potential impact from arrearages and the impact of rate subsidy reductions are larger. Some have found that programs that target high arrearage customers have particularly larger impacts from utility NEBs.

	Utility NEBs
	Few other impacts have been examined in great detail. If capacity impacts are examined and valued, it is likely peak programs will begin to have much more influential effects on Utility NEBs. To the extent line losses are higher or lower proportionally in peak vs. non-peak times; similar patterns will emerge if these values are incorporated.
Variations for Low Income or other sectors	Low income programs bring more Utility NEBs for arrearage reduction and reduced rate subsidies.
Variations by region of the country	Climate zones could affect these NEBs because of the effect of harsh winter climates (and high summer conditioning) on bills and arrearages, including for low income households. No specific patterns have been uncovered. In addition, gas utilities may see higher effects from potential emergency situations avoided.

Figure C.3: Patterns in Emissions and Job Impact NEBs by Type of Program and Region[112] (Source: Skumatz 2010)

	GHG Emissions	Economic Impacts
General results	Emissions impacts have improved a great deal over the last 5 years, and have shown significant impacts.	Range from multiplier of 3.54 for national expenditures on EE (Mulholland, Laitner, and Dietsch 2004) to multipliers of 0.25 for appliance replacement programs (Imbierowicz et. al. 2006). In OR one MW saved increases output by $2.2 million.
Variations by Program type	The effects vary significantly with program type to the extent that different programs deliver savings at different types of day / days of week / months of year. Emissions vary with the generation profile for the time the savings are delivered. Work by multiple authors finds these variations. Emissions reduction during peak hours is often smaller than for baseload reductions (baseload plants are less expensive but put off more GHG). However, see notes regarding region of country below. Thus air conditioner programs will have different profiles than lighting retrofits.	Dramatic impacts depending on program type because it affects different underlying industries affected by the program's specific measures and make-up (e.g. labor intensity). One study found multipliers from 30% more to more than doubled for weatherization compared to. Appliance replacement programs.[113] (Imbierowicz et. al 2006). The study finds that appliance replacement programs do not provide much multiplier effect even when national scope is considered, largely because appliances are mostly manufactured overseas
Variations by sector	No additional variations than by program type or region as listed elsewhere.	No additional variations than by program type or region as listed elsewhere.
Variations by region of the country	Significant variations by region of the country because the driver is electricity generation mix (at peak and off-peak). Where there is more hydro, emissions are lower, etc.	Variations are significant because the industry mix varies across the nation. The one study examining this impact[114] found that multiplier impacts for both weatherization and appliance replacement programs were always lower in Wisconsin than in California or nationwide (about 10% to 50% lower depending on program type). The study found slightly larger multipliers for California programs (likely due to broader industry mix), and largest when nationwide scope is considered.

Figure C.4: Variations in Participant NEBs by Program Type and Region (Source: Skumatz 2010)

---

[112] Again, note that California embeds emissions and T&D effects into the computations of avoided cost; no separate work on these NEBs is required. However, this summarizes the broader literature, for the interest of the reader, and the results may provide a value that can be compared to the values incorporated into the avoided cost.

[113] The study found economic output multipliers associated with weatherization program expenditures are considerably higher locally (more labor intensive) than those associated with appliance replacement programs (46% vs. 25% for WI, 49% vs. 34% for CA, and 106% vs. 25% US). (Imbierowicz, Skumatz, and Gardner 2006).

[114] Imbierowicz, Skumatz, and Gardner (2006)

	Participant NEBs
General results	Large – often equal to the value of the energy savings, depending on program (see below). There are patterns in leading NEBs as listed above.
Variations by Program type	Participant NEBs are higher for whole building programs than individual measure programs. This seems largely related to the inclusion of measures that affect comfort (HVAC, windows, design features).
Variations by sector	High value residential side NEBs tend to be: comfort, doing good for the environment, operations and maintenance / lifetime, and aesthetic effects. On the non-residential side, the most valued NEBs tend to relate to: comfort, operations and maintenance / lifetime, equipment performance, doing good for the environment, and labor / productivity issues. Low income programs tend to have higher NEB values associated with feature like "improved understanding of equipment energy use", control over bills, and similar. Negative NEBs – reflecting barriers – have also been measured. On the non-residential side, maintenance issues are the most common concern; on the residential side maintenance and aesthetic issues arise.
Variations by region of the country	Climate zones are influential in the value of NEBs because much of the high-value benefits come from comfort (affected by harsh winter climates and high summer conditioning). This single factor is often 15% or more of all participant NEBs. One study found that the highest valued source of NEBs was the insulation work (related to comfort).[115] In addition, on bills and arrearages, including for low income households. No specific patterns have been uncovered.

We can also examine the patterns by size and variability of NEB. Based on this analysis, the results show that – if a utility wanted to estimate the minimum of NEBs to minimize costs – the NEBs in the yellow cell (or potentially the pink cell) could be aggregated into a multiplier. The NEBs in the salmon or purple cells (high variation) either need further investigation to identify the source of variability (and thus, potentially turn them into multipliers, or adders based on those causal factors), or require estimation into the future because they are 1) important / highly valued, and/or 2) very program-specific. Not otherwise classified NEBs have not shown a clear pattern in value or variability.

[115] *Skumatz and Gardner 2004 decomposition study.*

Figure C.5: Variability and Patterns in Low Income NEBs (Source: Skumatz 2010)

	Large size NEB	Not elsewhere classified	Small size NEB
Low variation	None identified with this pattern		Arrearage and coll'n NEBs (but easily measured by program; also varies depending on whether target is "high arrearage" customers)
Not elsewhere classified		Insurance Substation / infrastructure Power quality Tax effects Health & Safety Wastewater / water infrastructure Social indicators T&D losses	
High variation	Emissions (predictable models) Economic impact (predictable models; depends on measures) Participant NEBs (depends on measures, household characteristics) Emergency gas service call (needs more analysis)		None identified with this pattern

APPENDIX D: STATE OF THE ART – AND GAPS – IN MEASUREMENT OF NEBS

This appendix is excerpted from Skumatz, "Non-Energy Benefits: Status, Findings, Next Steps, and Implications for Low Income Program Analyses in California", May 2010. The study addresses conclusions to low income programs, but the summary is applicable beyond this sector. See the original paper for the references.

Basic best-practices of NEBs have been fairly-well adopted within the literature. These include basics like including positive and negative NEBs, and consideration of "attributable" NEBs above what would have happened without the program. This last element assumes consideration of net-to-gross ratios; however, the special case of low income programs may support an assumption that the NTG is 1 because, in many cases, the investment may not have occurred without the program. Best practices include considering the following.

Categorization, Causes, and Uses of NEBs

Starting with work in the mid-1990s, the literature began to explore more consistent measurement methods, and sort these benefits into three "perspectives" based on the beneficiary of the effect – utility/agency; society, and participant.[116]

Considerations for Appropriate Attribution of NEB Impacts

The following is a list of basic issues to be considered in assessing and attributing NEB effects to EE interventions:

- **Redundancy in sources or categories**: Similarly-named benefits can arise in multiple perspectives without being redundant. For example, fewer billing-related calls to a utility save money and time for both the utility and the household making the call. These are distinct impacts. Of course, each needs to be valued in terms appropriate to that beneficiary, and the number of subsets of different perspectives and benefit categories that are included in a computation depends on what is appropriate for that specific application (e.g. particular benefit-cost tests, etc.).

- **"Net" Effects**: NEBS may be positive or negative, and the "net" effects may also be positive or negative. Negative benefits can be interpreted as barriers in some applications.

- **"Net" of standard equipment choices**: When NEBs are applied to energy efficiency programs, it is critical that the impact be measured above and beyond the base of what would happen without the program–specifically, the (presumably, standard efficiency) equipment that would be selected without the program.

- **"Net" of free riders**: To the extent that the interest is in NEBs that are attributable to the program above and beyond what would have happened without the intervention, the NEBs would have a free ridership (and potentially spillover) factor applied.

[116] *Initiated in Skumatz 1997 and described in detail in subsequent research,and repeated in Amann, 2006.*

- **Minimizing Overlap/Double Counting**: The drivers for NEB effects tend to emanate from a limited number of key impacts associated with energy efficient equipment. Multiple, closely related benefits and impacts could be measured, but it is likely the individual benefits might be difficult for participants to separately measure or assign value to each

State of Measurement and Results

The state of measurement of NEBs falls into several major categories. The traditional treatment, and concerns / revised considerations are discussed below.

Arrearage analyses: Arrearage studies for low income programs have been conducted for several decades, and are generally conducted using control and program groups, with straightforward analyses of the net impact of the (low income) program on arrearage, bad debt, consumer calls, shut-offs and reconnects, and other financial or "collections"-type factors. The statistical methods are well-known. There are scores of examples of these studies for utilities across the nation.

"Readily-measured NEBs": These NEBs are easily measured with direct computations of impacts or direct application of readily-accepted secondary data. An example of these computations includes the water savings from low flow showerheads or faucet aerators, or from efficient clothes washers, as well as the associated "soap" savings from these washers. These NEBs are computed based on average showers or laundry loads per household from established sources like the AWWA (American Water Works Association), or others, and the results tend to lead to minimum controversy.[117] These types of NEBs are measured around the country, but are formally included particularly in the Northwest, and are included for programs above and beyond just low income programs (particularly commercial / industrial programs).

Model-based societal NEBs: Third party models have been developed that provide well-founded estimates of the impacts of low income (and other) programs on emissions and on job creation / economic development.[118] These models are of varying degrees of detail / sophistication / cost, but the number of studies and models addressing these impacts (developed / published by universities and consultants) at the local, state, and national level are increasing – and are being accepted in the literature.

Survey-based Participant NEBs: Organized, statistical surveys have been used as the basis for computing a subset of participant-based NEBs since 1994. From nearly the beginning, the methods have been based on approaches drawn from the academic literature. The survey-based approaches have been used to measure the benefits related to: performance (comfort, etc.), lifetime, maintenance, property value, noise, safety, mobility, education impacts, "doing good" for the environment, and stability-type metrics, and any negative impacts associated with the programs. A number of main measurement approaches have been used for these survey-based studies: contingent valuation and willingness to pay / willingness to accept; relative scaling (percentage and labeled magnitude scaling); and ranking methods. Each has demonstrated academic and statistical underpinnings. The survey-based approach has been used for several reasons:

---

[117] Savings in other fuels may also be a potential category of NEBs that could be "readily measured".
[118] Note that the tax impacts of the economic development impacts have not been frequently measured, but would be fairly readily measured as well, given information on local tax codes.

- **Some of the values can <u>only</u> be derived from user perceptions**: Examples include: impacts related to knowledge / understanding of bills, feelings of doing good for the environment. It might be argued that perceptions of comfort are more relevant *<to participation decision-making, and to the participant's experience of comfort>* than measurements of thermal comfort.
- **Some of the values are most readily *<and are potentially most properly>* derived from user perceptions**, although they could theoretically be measured in other ways. Examples include: noise, thermal comfort, likelihood of moving due to high bills. In some cases studies are lacking that could provide independent[119] values for some program-related changes (e.g. sick days from work or from school, incidences of moving, etc.). In other cases, the studies to conduct the analyses on a program-by-program basis would be expensive[120] (e.g. metering statistical samples of homes for noise, lumens, temperatures), or if the incidences of occurrences are low and would require many samples to identify impacts (for example, high value health and safety events).
- **Surveys are the fastest way of gathering data on multiple NEB categories**. This is certainly true; however, the values gathered via survey should be compared with the values computed via other means to assess the credibility and consistency of survey-based measures.[121]

Based on further analysis, we believe some of the NEB categories that have been measured via survey could and should be moved from survey-based estimation methods to more direct financial computations / estimations (see next category).

Financial Computations: The potential exists to use age, manufacture data, and third party information to compute some NEB values in low income programs; however, this has rarely (or never, as far as we can find) been done. The most appropriate NEBs for this approach include valuations from lifetimes or from maintenance. Using information on the average age (cohorts) of equipment replaced in the participant homes (to be gathered as part of program records) and records / expectations related to new equipment, replicable valuations for these types of NEBs could be computed.

Weak / unexplored NEBs: A number of NEBs have barely or never been measured. These include, most particularly, a host of important health-and-safety effects relating to both the participants and utility, including utility insurance savings; indoor air quality impacts (particularly on occupant health); doctor visits, etc. A number of others have also been little-explored, including national security, tax benefits, and others.

There are several additional notable measurement issues in NEBs in addition to those discussed above:

1) Statistical / academic grounding: There are several threads of the survey-based NEB literature that specifically address the statistical and academic grounding for the use of the survey method(s). These include: work by Skumatz or Skumatz and Gardner (about a dozen papers starting in 1995); a paper by Summit Blue (2007) and several papers by Lutzenhiser.

---

[119] *and potentially transferable, at least within climate zones*
[120] *For some it would be expensive relative to the potential values, although this needs to be better demonstrated*
[121] *Literature has suggested that for businesses, specific research on key topics by those businesses may be a valuable and especially accurate source of information on the measure's NEBs. However, 1) that is not very practical for low income programs, and 2) the statistical reliability of those estimates in a commercial setting are suspect, as only a few businesses would be conducting these studies, and those results would tend to be computed only for businesses that did, or expected to have, large values for that NEB, biasing the ultimate results.*

2) Use of regression analysis for estimating impacts: Researchers at Heschong-Mahone used regression approaches to relate academic test scores to daylighting in schools, and sales to daylighting in retail outlets. However, these methods have not been applied to low income programs or measures, and show most promise for measuring just a couple of NEB effects at a time, and require considerable data collection to control for other contributing factors (affecting, for example, sales or test scores).

3) Comparisons of values derived from different survey measurement methods: Two authors have conducted the bulk of this type of work: Skumatz (many papers, starting in 2000) and Hall (2007). More work of this nature is important to identify the most credible, consistent, and robust measurement methods.

4) Cross-program studies identifying patterns in NEBs (sizes and variability): Few studies have looked beyond the single utility program being analyzed to compare results to other programs. The exceptions for low income include: Skumatz (1998 and others), Hall et.al.(2007), Skumatz and Cadmus (2009).[122]

5) Measure- vs. Program- Based NEBs: Within the low income sector, almost all NEB work has been conducted as program-wide estimations.[123] One study (Skumatz and Gardner 2004) *<used regression analysis to successfully>* test the potential of disaggregating program-wide NEBs to the specific measures installed. Although NEBs from appliances have been measured, measure-based NEB work has not been conducted estimating NEBs from insulation, caulking, education, or many of the types of measures included in California- and other low income programs.[124]

[122] Later note: Amann 2006 compares results from seven programs; five were SERA studies.
[123] Measure-based work has been conducted for commercial – industrial programs (which tend to be measure- based, like boiler, motor, and lighting studies).
[124] Some household appliances have had specific NEB estimation work, including clothes washers, air conditioners, refrigerators, dish washers, and CFLs. Skumatz has conducted some work on just insulation, but this is related to measures installed overseas, not in US low income programs.