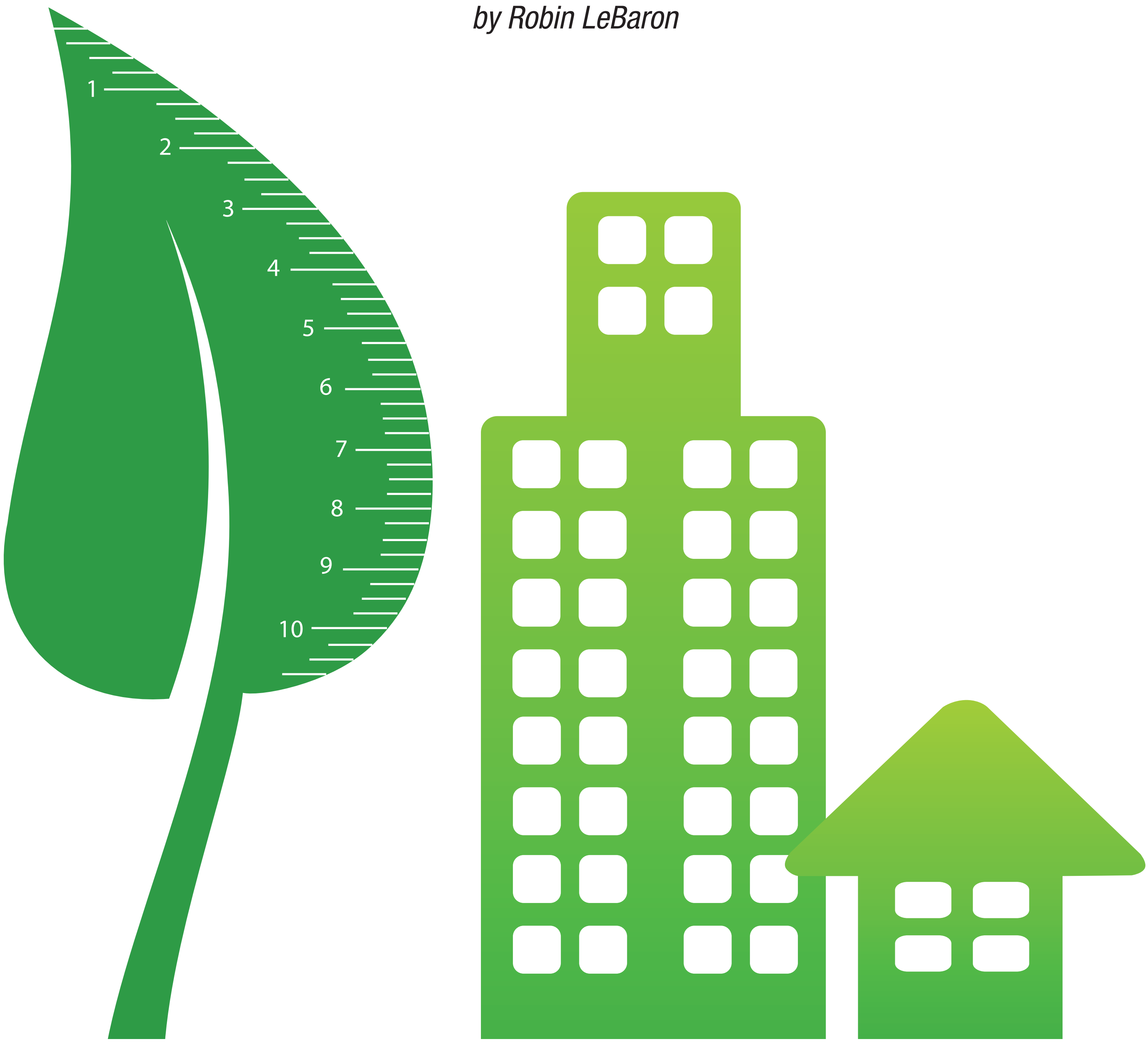
National Home
Performance Council

June 2012

MEASURE IT RIGHT

Best Practices in the Selection and Implementation of Cost-Effectiveness Tests

by Robin LeBaron



Foreword

The National Home Performance Council is a national non-profit organization created to support whole-house energy efficiency programs through research and stakeholder engagement. NHPC is stakeholder-driven, with a board of directors that includes a wide range of energy efficiency stakeholders: Federal agencies, state energy offices, non-profit organizations, contractors, program implementers, real estate representatives, utilities, and manufacturers. NHPC's mission is to address challenges that prevent the growth and expansion of the whole-house energy efficiency sector.

In the spring of 2011, a group of stakeholders convened to explore how cost-effectiveness tests were affecting whole-house energy efficiency programs. It was agreed that further research on the issue was needed. With the support of its board, NHPC assumed responsibility for coordinating the stakeholder group, and for conducting research that would provide preliminary answers to the group's questions.

This paper is the outcome of this effort. It identifies and describes problems raised by stakeholders, and proposes methods to address these problems in both the long and short-term. Its final recommendation is further research: but research, in this case, on test implementation "best practices" that could be put to immediate and practical use that would benefit all energy efficiency program stakeholders.

Introduction

The suite of tests used to test the cost-effectiveness of energy efficiency and other programs funded by ratepayer charges are creating significant challenges for energy efficiency programs. These tests have an important rationale: to ensure that ratepayer funds are spent on programs that will provide genuine benefits to ratepayers and other stakeholders. As currently structured and implemented in many jurisdictions, however, the tests can constrain the design and implementation of residential energy efficiency programs, particularly programs intended to support comprehensive energy efficiency upgrades. As a result, cost-effectiveness tests may in effect work at cross purposes to important public policy goals, such as carbon reduction, clean air, and job creation, and even national security and reduction in dependence on foreign energy sources.

Energy efficiency practitioners have voiced concerns about the cost-effectiveness tests, and proposed changes to the testing methodologies, for decades. Their concerns include the fact that implementation of the tests has prevented the creation of good programs in some jurisdictions, and has shaped the design of existing programs in ways that reduce their effectiveness, restrict consumer choice, and undercut a range of other public policy goals. In the past, when whole-house retrofit programs were relatively new, and single-measure programs offered real savings, the limitations of the cost-effectiveness tests and the problems they caused could be overlooked or deferred. Increasingly, however, utilities are required to achieve energy savings targets. Whole house upgrade and other programs that offer “deeper” energy savings have matured and can make significant contributions to meeting these goals. As a result, legislatures, commissions, utilities and program administrators need tests that give the best and fullest answers regarding which programs should be supported with public and ratepayer dollars. Some methods of test implementation, however, may limit the information that these stakeholders have available, possibly preventing the realization of significant energy efficiency gains.

This paper describes the problems and issues that arise for energy efficiency programs, particularly whole-house programs, as a result of common cost-effectiveness test implementation practices. To address these challenges, the paper recommends:

1. That the Societal Cost Test (or, as a less preferable alternative, the Total Resource Cost Test) be used as the primary test for program evaluation, but only if the test can be conducted according to best practices that include appropriate consideration of the full range of costs and benefits.
2. That if it is impossible or cost-prohibitive to implement the SCT (or TRC) according to best practices, the PAC should be used as the primary test.

The Cost-Effectiveness Tests

In simple terms, a cost-effectiveness test is designed to determine whether the benefits of a particular program (or, as discussed below, portfolio, project or measure) outweigh its costs. These tests can be expressed as a net present value or as a ratio of benefits to costs. In the second case, a result of greater than one indicates that benefits outweigh costs, while a result of less than one indicates the reverse. In theory, the exercise sounds simple; in practice, the results depend on the details of how "costs" and "benefits" are defined.

For investor-owned utilities, the basic decisions regarding which tests should be used for evaluative purposes and how they should be administered are most typically made by legislative bodies or state-level commissions (commonly known as public service or public utility commissions). With only a few exceptions, each state uses one, or a combination, of five tests developed in the early 1980s by the California Public Utilities Commission and codified in the California Standard Practices Manual. Each of these tests is designed to compare the benefits and costs of a ratepayer-funded program from the perspective of a particular stakeholders: a program participant (the Participant Cost test), non-participant ratepayers (the Ratepayer Impact Measure), the program administrator (the Program Administrator Cost test), the combination of utility and ratepayers (both participants and non-participants), within a utility service area (the Total Resource Cost test) and the society served by the utility in question (the Societal Cost test).[1]

The Participant Cost Test: The Participant Cost test ensures that a participant in the program will benefit, i.e., that the savings realized by the participant will be greater than the costs that she or he incurs. This is an important minimum standard, but commissions and utilities will typically seek to conduct other tests to ensure that the program is cost-effective from the perspective of other stakeholders. In other words, a program should clear the Participant Cost test, but the test should not be the only measure of a program's cost-effectiveness.

The Ratepayer Impact Measure Test: The RIM is designed to assess the impact that an energy efficiency program will have on utility rates. The RIM is typically used to determine whether non-participants in a program, i.e. utility customers who do not take advantage of the energy efficiency program under consideration, will be affected by higher rates as a result of the program. Rates are determined on the basis of a per-unit charge for energy that allows a utility to earn a given return on equity. If the amount of energy provided by the utility falls, it may be allowed to raise rates to make up for the shortfall. Because energy efficiency programs are designed to decrease energy consumption, they almost by definition result in rate increases in the medium term, as per the RIM.

However, the RIM provides no real indication of the magnitude of the rate impact or its timing. Even more significantly, the RIM often does not take into account the potential for energy efficiency measures to defer new capital investments in capacity or distribution that would significantly increase rates. As a result of the RIM's limitations, a program could fail the test even if it created large savings for participants, while resulting

1 As excellent descriptions of these tests are available in a number of existing sources, this paper does not provide detailed descriptions; readers interested in learning more may want to consult National Action Plan for Energy Efficiency 2008 for an introduction to these issues.

in only minimal costs (in the form of bill increases) for non-participants. For these reasons, the RIM does not provide a meaningful basis for making determinations as to whether an energy efficiency program should be funded, and should not be used for this purpose. As a recent national State and Local Energy Efficiency Action report observes: "The Ratepayer Impact Measure (RIM) test is an insufficient indicator of rate impacts, as it is overly narrow and does not present rate and bill impact information in a way that is useful to regulatory commissions," (SEE Action 2011: iii). As of early 2012, the last state to use the RIM test as its primary screening tool, Virginia, changed its system for screening programs in a way that results in the RIM no longer serving as the determining test.

The Total Resource and Societal Cost Tests: The Societal Cost Test and Total Resource Cost Test (TRC) are theoretically the most appropriate tests for evaluating energy efficiency programs. Both are designed to compare the costs and benefits of a program from the perspective all ratepayers in the utility service territory. The TRC is designed to determine whether a program increases or decreases the cost of energy within the utility's service territory. The SCT incorporates considerations of costs and benefits that are not quantifiable, thus providing the highest-level review of the impact of an energy efficiency program.

According to ACEEE, 29 states currently use the TRC as their primary screening tool, and six states use the SCT.

The most significant limitation of the TRC and SCT is that the full range of costs and benefits need to be considered in the test, or the test results will fail to allow an accurate interpretation of the program's potential. Because the test inputs are complex, and can be time-consuming and expensive to measure, however, the tests may be implemented without information about the most difficult-to-measure, yet crucial, inputs. This can result in significantly skewed tests results, particularly because the benefits tend to be inherently more difficult to measure than the costs.

The issue of consumer motivation illustrates the practical challenges involved in implementing the TRC and SCT. Because both tests consider the full range of costs involved in a program, they include the amount that a participant contributes to the cost of installing energy efficiency measures as well as utility incentives. But the participant might have made some if not all of the investment on their own, even in the absence of the program. Similarly, the participant may derive a wide range of benefits from the energy efficiency measures installed in addition to energy savings: in some cases, non-energy-related considerations, notably increased comfort, may drive a consumer's decision to implement energy efficiency measures. To conduct a TRC or SCT test accurately, it is crucial to ensure that only costs that the consumer would not otherwise have spent (incremental costs) are included, and that other non-energy-related benefits that the consumer derives are taken into account. Because both factors are inherently difficult to measure, however, they are not always considered. Similar issues arise in determining the net-to-gross calculation of program benefits: the benefits may be reduced by an estimate of free ridership, but spillover and market transformation, which may seem more difficult to quantify, are less frequently included in test results.

Publicly funded programs frequently seek to encourage private investment, and measure success in part in terms of private dollars leveraged to achieve program goals. The TRC and SCT, by contrast, effectively penalize programs that successfully encourage consumer investment by including these costs in the cost-benefit calculation.

The SCT and TRC then, while the most appropriate tests in theory, may not be the best tests in practice. If a program cannot apply the SCT (or TRC) rigorously, including reasonable assessments of all relevant costs and benefits, it is not providing a full answer to the fundamental question posed by the test. In these cases, commissions should not use the SCT (or TRC) for making the final determination as to which programs should be supported with ratepayer funds.

The Program Administrator Test: The Program Administrator test measures whether the benefits to the utility or program administrator are greater than the costs that the utility incurs to run the program, including administrative costs and subsidies. In a sense, the test can be seen as one method for addressing the question of whether a program will provide energy at a lower cost than other, generally supply-side, options. Because the test does not include participant contributions, it involves fewer methodological challenges, and can be easier and less expensive to administer. However, the test can be limited in assessing comprehensive (whole house) programs when the sponsoring utility does not provide all the energy to the home, either because a separate utility does so, as in the case of jurisdictions served by separate gas and electric utilities, or because bulk fuels are common in the region.

Implications of Measuring it Wrong

The preceding discussion makes clear that the cost-effectiveness tests are complex in concept and execution. Selecting the correct test, and implementing it correctly, can present a commission with a significant challenge. What is at stake if the tests are not implemented appropriately?

Cost-effectiveness tests can have an extremely significant impact on energy efficiency programs. At the most basic level, the tests can prevent a program from being created. In Virginia, for example, use of the RIM test until very recently effectively blocked the creation of a wide range of energy efficiency programs proposed by utilities and advocates, because the programs all failed the test.

Less obviously, the tests sometimes shape programmatic design and implementation. Because customer costs are included in TRC and SCT calculations, measures that involve significant customer outlays, such as new high-efficiency HVAC equipment, can detract from a program's ability to pass the TRC, and may thus be excluded from the program as a result. In addition to reducing a program's ability to realize deeper energy savings, this can reduce the opportunity to realize benefits from interactions among measures, such as those that result from appropriate sizing of HVAC equipment following insulation and air-sealing work. The measures excluded from a program as a result of test considerations may also be significant drivers of consumer interest, thus potentially reducing the program's success and uptake.

Cost-effectiveness test considerations may shape a program in other ways. Testing of each project or measure, for example, may require the development of expensive programmatic infrastructure to conduct tests for each new job. Such requirements detract significantly from the ability of contractors working in market-based programs to close sales, because the testing required is time consuming, may confuse the customer, and may result in elimination of the measures that the customer is really interested in.

In extreme cases, cost-effectiveness tests may lead to the elimination or reduction of thriving programs. A specific instance of this more general program involves programs that were launched with ARRA funding and now need new sponsors: utilities may decide not to adopt such programs in part as a result of concerns about their ability to clear the cost-effectiveness tests.

In short, cost-effectiveness tests are creating obstacles to the creation and effective design of energy efficiency programs, particularly complex programs designed to achieve relatively deep energy savings. Without these programs, it will be difficult for states operating with energy efficiency portfolio standards to meet their savings goals. More importantly, if such programs are never created, they will never be able to realize their potential to limit rate increases over the long term by deferring the need for capital expenditures on generation, transmission and distribution capacity. To the extent that the cost-effectiveness tests prevent utilities and programs from pursuing such opportunities, they do a great disservice to ratepayers.

Measuring it Right: Best Practices, or PACT

Given the implications of getting the tests wrong, it is crucial that the correct test be used to evaluate energy efficiency measures, and that the test used be implemented correctly. This paper recommends identification and use of best practices in application of the tests. Best practices could address many of the problems discussed above, creating benefits for all stakeholders.

As noted above, in theory, the SCT and TRC are the most appropriate tests to use to determine whether an energy efficiency program will benefit ratepayers. As noted above, 35 states currently use either the SCT (six states) or TRC (29 states) as the primary test for evaluating energy efficiency programs. Because the California Standard Practice manual allows considerable latitude in precisely how these tests are conducted, test implementation varies tremendously from state to state, both in the choice of tests and in the methodology used to implement them. Although some of these differences may reflect specific local conditions, many others result from ad hoc or legacy decision-making about test implementation that may not be fully consistent with the tests' underlying rationale or the current interest of the commission and ratepayers.

The SCT and TRC are variations on a single test methodology designed to measure the costs and benefits of a program from a holistic perspective that includes both program participants and non-participants in the utility service territory. In essence, the TRC provides an indication as to whether the total cost of energy will increase or decrease as a result of the program, while the SCT asks a similar question but includes externalities that are not factored into existing price mechanisms. The SCT thus provides the fullest perspective on an efficiency program, by taking into account the full range of program impacts, while the TRC in theory looks only at the somewhat more narrow issue of quantifiable costs.

Because the SCT offers most comprehensive perspective on the costs and benefits of a program, it is recommended as the most appropriate test for screening energy efficiency programs. One of the reasons that the SCT is not used more frequently is that the inputs are difficult to measure. However, in practice, this problem applies to the TRC as well, as neither test is really applied appropriately unless all costs and benefits are considered – and some of these are inherently difficult to quantify. In other words, the dividing line between the two tests is less clear in practice than in theory, and given that a certain arbitrariness may result in deciding what should and should not be included in the TRC, it is more rational to seek to compare all costs and benefits. However, the TRC is also appropriate for program screening purposes if a commission or legislature believes that the SCT is not appropriate for their specific situation.

Whether applying the SCT or the TRC, it is important to ensure that the tests are applied in a way that adheres to the underlying principles and goals of the cost-effectiveness tests. As noted before, the SCT and TRC can be seen as designed, respectively, to determine whether a program provides a net benefit to society, and whether the program lowers the cost of energy within a service territory – although because the challenges involved in quantification of test inputs mean that in practice the line between these questions is less distinct than this summary statement suggests.

Best practices should be designed to support test methods that adhere to these underlying test goals. A test that systematically measures all benefits while ignoring a set of costs, or vice versa, is not really designed to answer the question set by the SCT and TRC tests. The first requirement of best practices, accordingly, is that they should be designed to ensure that test methods support realization of the test's underlying goals. A second requirement, which should complement rather than working at cross-purposes to the first, is that best practices should support the efficient and effective functioning of programs, i.e. they should not create unnecessary or arbitrary constraints on programs that are designed to achieve a public good.

The following section outlines a set of methods for applying the SCT and TRC that conform to the criteria for best practices outline above. This list is not intended to be definitive, but to promote discussion and elaboration of best practices that can achieve industry consensus. Because the SCT and TRC are variations on a single test methodology, most of the best practice recommendations that follow apply to both tests. The exception is the recommendation regarding the discount rate, which is relevant primarily for the TRC test.

Spillover and market transformation in the net-to-gross calculation

Proposed best practices:
Include spillover and market transformation effects as well as free-ridership in the net to gross adjustment, or use gross savings if no data on spillover and market transformation is available.

The net to gross adjustment to the TRC calculation is designed to address the fact that projected energy savings may be different than actual energy savings, resulting in net energy savings that are different than predicted (gross) savings. One of the main challenges in the net-gross calculation is determining the extent of free-ridership, which should affect the calculation such that gross savings are reduced by the amount that program participants would have expended even in the absence of subsidies. In practice, however, the methodological challenges involved in estimating free ridership are significant. Moreover, whole-house programs are designed to (and do) encourage spillover and market transformation; for example, in cases in which participants decide to implement additional (non-subsidized) measures, or non-participants decide to implement energy efficiency measures with their own resources without participating in the program. These problems highlight an inconsistency in many applications of the SCT and TRC: they often measure total costs (including costs that should logically be excluded), but are explicitly designed to capture only net savings (Eckman 7). One way to address this issue is to factor in spillover and market transformation effects, as well as to ensure that only net costs are being measured.

Some New England jurisdictions assume that spillover cancels out free ridership, and hence just use gross savings, although the New England ISO has not approved this methodology (Titus and Michels 2008: 5-316).

Full recognize all avoided costs

The avoided costs that are not incurred as a result of the energy efficiency program are a crucial benefit in a cost-effectiveness test. Energy efficiency programs have the capacity to defer a number of types of costs, including energy, capacity, and transmission and distribution (T&D) costs. Each type of cost can be measured in multiple ways, some of which rely on relatively simple assumptions, and some of which are extremely technical and detailed. Estimates of T&D cost savings, for example, can take into account the impact of energy efficiency savings on specific planned projects. Each of these costs should be incorporated into the cost-effectiveness testing process, and the cost estimates should strive for the highest level of accuracy possible, although cost constraints will be significant in some cases.

Proposed best practices: Include all avoided costs (e.g., T&D costs as well as energy and capacity costs), and use the most rigorous methods reasonable available to determine costs.

Effective Useful Life (EUL) of measures

The effective useful life (or EUL) of a measure is the span of time that it can be reasonably expected to function as intended. EULs range substantially, depending on the product and the manufacturer: a CFL might typically last between five and ten years, while a window might function for up to 75 years, and building insulation 100 years or more. Many programs, however, cap EULs at around 20 years (Hall et al. 2008), both because of the expectation that homeowners may replace measures prior to the end of their EUL, and because the discount rates frequently used by programs devalue savings that continue beyond 25 years. For the purposes of cost-effectiveness testing, this significantly reduces the value of measures with long useful lives. Such measures include a number that are significant for whole-house upgrades, notably insulation.

Proposed best practices: Programs should not use arbitrary caps on EUL.EULs may be adjusted for the possibility of consumer replacement, but such adjustments should not be excessive, and should be based on research on consumer behavior to the greatest extent possible.

Time frame

Applications of the TRC may consider the costs and benefits associated with a program over the span of one year. This approach does not take into account the fact that a program's cost structure is likely to change over time. As long-standing home performance programs, such as the one sponsored by NYSERDA in New York, have demonstrated, whole-house energy efficiency initiatives require large up-front investments and may only achieve their full impact several years into the program's life cycle, once they have begun to transform local markets. Taking into account average costs and benefits over a multi-year time frame is one way to address this issue.

Proposed best practices: If a program has significant up-front costs, the SCT/TRC should amortize the costs over multiple years. An alternative approach would be to conduct an evaluation of the program over multiple years if the testing process allows this possibility.

Recognize all energy savings and distribute costs appropriately

In areas served by separate electric and gas utilities in which a program is sponsored by only one of the two utilities (most frequently the electric utility), the SCT and TRC may fail to fully account savings in other fuels (i.e., fuels not provided by the sponsor utility). Bulk fuels in particular may be missed by the tests. Because reduction on consumption of all fuels results in real, quantifiable cost savings, they should be incorporated into the SCT and TRC as a matter of course.

Other methodological shortcomings may also occur in dual fuel jurisdictions. If a program or measured is incentivized by two utilities, costs as well as benefits should be apportioned appropriately, so that neither is double counted. Although this seems obvious in theory, in some cases costs in particular may be double counted; i.e. applied for each program.

Proposed best practices: Consider all energy savings, not just those obtained by the participating utility, and ensure that costs are not double counted.

Proposed best practices: For whole-house programs, discount participant costs by a percentage that reflects the average value of non-energy benefits relative to energy savings, using national or regional average values. For single-measure programs, reduce participant costs by the estimated value of non-energy benefits. Non-energy costs should be incorporated into the tests to the extent that there is a basis to believe they are significant.[2]

Non-energy impacts

Non-energy impacts (NEIs) are the impacts other than energy savings that result from the implementation of an energy efficiency program. Many of these impacts are non-energy benefits (NEBs). For homeowners, these may include comfort, health and safety, aesthetics, and other general quality of life issues, as well as financial savings such that result from causes other than increases in energy efficiency, such as savings on water costs resulting from the installation of a high-efficiency dishwasher or washing machine. These benefits can be substantial: an Oak Ridge National Laboratory study found that non-energy benefits such as comfort slightly outweighed the value of energy benefits for Weatherization program clients (see Schweitzer and Tonn, 2002). The data suggests that these benefits, particularly comfort, are frequently the primary motivator for consumers who implement energy efficiency upgrades. Other social non-energy benefits include water savings, local job creation and economic development: these too can be significant.

Methods have been developed to quantify the value of these benefits, but they are generally complex and expensive to administer, and there is no widespread consensus regarding which are the most appropriate. As a result, many states have not incorporated NEIs into the TRC, and those that do tend to use conservative estimates for valuing NEBs (Amann 2006: iii). The evidence for the significance of NEBs is at this point significant enough that they should be taken into account by any SCT and TRC tests, on the grounds that without consideration of NEBs the test is effectively asymmetrical: it captures the full value of participant costs, but ignores extremely significant benefits that are real drivers for program participants.

In theory, the value of non-energy benefits can be considered a benefit for the purposes of applying the SCT/TRC test. Given that these benefits fall outside the scope of the utility commission's jurisdiction, however, a more appropriate methodology is to reduce participant costs by the estimated value of NEBs. A simpler method is to determine the average value of NEBs associated with a program and discount participant costs by a percentage that reflects the ratio of the average NEB value to energy savings. Average NEB value could be determined at a national or regional level through well-designed survey research in a cost-effective fashion for use in multiple SCT/TRC tests. (See Amann 2006: 13.)

It should be noted that programs may also create non-energy costs, particularly in the form of lost time or lower satisfaction resulting from the increased complexity or other design features of energy efficiency measures. To the extent that there is plausible reason to believe that such costs are significant, they should be treated in the same manner as non-energy benefits.

It should be noted that programs may create non-energy costs as well as non-energy benefits. To achieve an accurate test result, non-energy costs should also be included in tests whenever relevant.

---

2 The average discount approach is methodologicaly simple, and it works for whole-house programs. For single-measure approaches, valuing the individual measures could be a more effective approach.

Future avoided costs of environmental regulation

Proposed best practices: Include reasonable assessments of avoided costs of all probable future environmental regulation.

The cost of complying with current environmental regulations is now frequently included in the SCT and TRC as one of the avoided costs of energy. Reasonably expected costs for future environmental regulations, by contrast, are rarely if ever considered. Given that many such costs, notably those associated with recent U.S. Environmental Protection Agency (EPA) regulations, are almost certain to be incurred in the future and are amenable to quantification, they should logically be included in both the SCT and the TRC.

To the extent that carbon emissions are or will be regulated in the future, for example through EPA's or through regional greenhouse gas initiatives.

It should be noted that there is a strong case for inclusion of consideration of other environmental externalities that have a direct impact on persons living within the utility's service territory but do not result in costs to the utility. To the extent that these represent avoided impacts, they may be more appropriately considered non-energy impacts.

Discount rate

Proposed best practices: Use a societal discount rate as a ceiling for SCT and TRC testing. Alternately, if the TRC is the test, use a WACC that reflects the relatively low risk involved in the creation of energy efficiency resources. Jurisdictions may consider using an even lower rate to reflect the increasing value of avoided carbon emissions over time.

In calculating the SCT and TRC tests, costs and benefits are presented in terms of their net present value (NPV) – their value taking into account the opportunity cost of money, or discount rate. Because costs are incurred up front, while savings are realized over an extended period of time, higher discount rates typically result in the program scoring lower on the test. The discount rate can have a substantial impact on test outcomes, and also decrease the value of the effective useful life (EUL) of measures with long life-spans, as discussed below.

The discount rate to be used depends on the specific test and stakeholder interests. The Participant test, for example, should theoretically use a rate equivalent to an average of consumer loan rates for similar products. For the Societal test, a "societal" rate comprised of a blend Treasury bill rates is commonly employed. The discount rate used in the TRC test varies considerably; some states use a societal discount rate, while others employ the weighted average cost of capital (WACC) for the utility.

As discussed above, this paper recommends use of the Societal Cost Test (SCT) on the grounds that it is the most comprehensive test and is best designed to inform public policy decisions. If the TRC is used, the choice of discount rate is ambiguous, given the composite, "society-like" nature of the stakeholder considered by the test. If WACC is used, consideration should be given to the fact that energy efficiency is a low-risk

Proposed best practices: Utilities should conduct the SCT (and TRC) tests at the portfolio level, and should have full latitude to retain programs that do not pass the SCT or TRC but that address other compelling policy objectives if the overall energy efficiency portfolio is cost effective. The SCT (TRC) test can be applied to individual projects or measures for informational purposes, but these results should not disqualify projects or measure if the program in which they are incorporated passes the SCT (TRC).

resource. A societal discount rate is a more obvious choice, however, even for the TRC, because the test is looking at a composite subject that is, in effect, "society," and because the programs under consideration inevitably have significant public policy implications.

Level of application

The SCT and TRC tests can be used to determine the cost-effectiveness of different levels of a program administrator's portfolio. At the broadest level, the test can be applied only to the entire portfolio of energy efficiency programs, allowing the highly efficient programs to "support" less efficient ones. This approach provides the program administrator with some latitude to experiment, bring new efforts on line, and sustain programs that target hard-to-reach market segments, while still ensuring that the overall suite of programs represents an effective use of ratepayer dollars. By contrast, the test is sometimes applied to individual energy conservation measures, which severely curtails the scope of a potential whole-house retrofit program by eliminating the possibility of supporting measures with long pay-back periods. Requiring application of the test at this level for each job can burden a program with administrative tasks that drive up costs for both the program implementer and contractors, while making it less attractive to consumers. It also burdens the installation contractors and their potential customers with a cumbersome process at the point of sale, making it more difficult to close sales. Applying the SCT or TRC test at the project level (i.e., the level of a house in the case of a whole-house program) is less restrictive than application at the measure level, but still interferes with the potential for achieving greater energy savings as the result of a comprehensive project. Application at the program level is even less restrictive, although it still does not provide program administrators with the flexibility of a test administered at the portfolio level.

Applied together, these best practices address the many logical inconsistencies that currently characterize many applications of the SCT/TRC test, and allow for standardized comparisons of test results across jurisdictions.

Use the Program Administrator Cost Test When the SCT or TRC Cannot Be Implemented According to Best Practices

As the foregoing discussion illustrates, the SCT and TRC require many inputs that can be complex, difficult and expensive to administer. Accordingly, the second recommendation is that in cases in which a commission determines that it is not possible or cost-effective to gather the information necessary to implement the SCT or TRC according to best practices, use of the Program Administrator Cost (PAC) test is recommended as the tool for determining whether an energy efficiency program should be supported with ratepayer funds.

The Program Administrator Cost test attempts to compare what the program in question costs the program administrator, whether a utility or other administrative entity, to what it would cost the administrator to generate the same energy from the least expensive supply-side source. Stated another way, the test compares the cost of conserving energy to the cost of supplying the equivalent amount of energy. This approach does not provide the same global perspective as the SCT or TRC. However, it does offer three significant benefits when contrasted to the TRC and SCT.

First, the PAC test is relatively simple to administer because, unlike the SCT and TRC, it does not require quantification of costs and benefits that are inherently difficult to quantify. The PAC considers the cost of program administration and subsidies, both of which are easy to measure. Measurement of non-energy benefits and the incremental cost of high-efficiency measures, which are an important part of the TRC test, can be avoided. As a collateral benefit, eliminating the need to measure these factors can result in substantial savings for a program, as the measurement processes are often expensive.

Second, the PAC test takes a preliminary step towards addressing the asymmetry between supply and demand-side resources that the cost-effectiveness tests impose. Applying the SCT or TRC as a screening tool creates an artificial barrier for demand-side resources (such as energy efficiency) that supply-side resources, traditional or otherwise, do not have to surmount. This is because demand-side resources must demonstrate that they do not impose a net cost to society, while supply-side resources are judged only by their price relative to the prices of other potential sources of energy. As Neme and Kushler summarize the issue:

> *Many observers have pointed out that other supply-side power sources are not evaluated according the cost of production, but only by the cost to the utility. Even when the price of a supply-side resource reflects substantial public subsidies, the cost factors involved in generating the resource are irrelevant to the utility's decision to purchase supply-side resources. (Neme and Kushler 2010: 5-304)*

This issue is important because large public subsidies are frequently deployed in support of supply-side resources. The California Standard Practice Manual recognizes this issue in the observation that "[s] upply-side resource options are typically based only on the costs incurred by the power suppliers," (CSPM 2001:21), but does not draw the conclusions that logically follow; namely, that use of the TRC as a primary screening measure results in an inherently unfair comparison.

The third benefit of the PAC test when compared to the SCT (or TRC) test is that the PAC allows consumer expenditures on energy efficiency to be seen as a positive outcome rather than as a cost. While in theory, from a societal perspective consumer spending on energy efficient measures is an expense, it is an expense that the consumer willingly and knowingly assumes, presumably because she or he expects to benefit from it. At a time of limited public resources, programs that successfully leverage consumer spending to meet both public and private goals should be encouraged rather than penalized.

A significant limitation of the PAC test, by contrast, is that it does not provide an effective tool for measuring a whole-house program in an area served by separate gas and electric utilities, for the obvious reason that if only a single utility participates it will only benefit from reductions in use of the fuel that it provides. In dual fuel markets in which only one utility sponsors a program, use of the PAC is probably not an option.

Conclusions and Next Steps

Energy efficiency is the quickest, cleanest, and cheapest means of meeting U.S. energy needs. To fully realize the potential of energy efficiency, it is important that the cost-effectiveness tests used to evaluate energy efficiency programs serve as tools to analyze and ensure effective implementation. As discussed above, present application of the tests inadvertently creates unnecessary barriers to the creation and development of good programs in many states. To help address these challenges, this paper proposes use of the most comprehensive of the cost-effectiveness tests, the Societal Cost Test or Total Resource Cost Test, and recommends development of best practices to ensure that the SCT or the TRC are applied in a way consistent with the underlying rationale of the tests.

The list of proposed best practices presented in this paper is designed to encourage further discussion and analysis of the issue. Further, more in-depth research on the best practice methodologies for SCT and TRC implementation would provide commissions and other regulatory bodies with better and more consistent information to aid in decision-making and would address some of the problems and issues related to current test implementation. Stakeholder review of the research would help to further refine and improve the findings.

Further research should also be conducted on the related subject of rate and bill impacts. The bodies that oversee the cost-effectiveness tests –commissions, legislators, or others – are typically concerned that energy efficiency programs not have a harmful impact on ratepayer's energy costs. As discussed above, it is generally recognized that the RIM test does not adequately address this issue. In theory, a properly designed SCT or TRC test should ensure that programs achieve a net reduction in energy costs within a service territory. However, the subject deserves more in-depth treatment that would look at both bill and rate impacts over the short and long terms. Such research would complement work on best practices in test implementation by expanding the range of information available to commissioners and legislators to use when making decisions about energy efficiency programs.

Finally, it should be noted that the entire suite of cost-effectiveness tests has its limitations. In the end, the most appropriate treatment for energy efficiency is as a resource – an alternative strategy for meeting energy needs. None of the current cost-effectiveness tests provide such an evaluation: the Program Administrator Cost test (PAC) comes closest, but does not really provide the an adequate tool to compare the cost of creating "negawatts" of energy efficiency to the cost of traditional supply-side energy sources. Over the longer term, strategies to support such comparisons in ways that benefit all stakeholders –ratepayers and utilities – should be developed. Until this can be accomplished, it is important to ensure that the existing, widely-used tests are applied as judiciously and accurately as possible through the use of best practices.

References

Amann, J. 2006. “Valuation of Non-Energy Benefits to Determine Cost-Effectiveness of Whole-House Retrofits Programs: A Literature Review.” Washington, DC. American Council for an Energy-Efficient Economy. LINK

Cappers, P., C. Goldman, M. Chait, G. Edgar, J. Schlegel and W. Shirley. 2009. “Financial Analysis of Incentive Mechanisms to Promote Energy Efficiency: Case Study of a Prototypical Southwest Utility.” LBNL Report Summary. LINK

CPUC (California Energy Commission & California Public Utilities Commission). 2001. “California Standard Practice Manual – Economic Analysis Of Demand Side Programs And Projects.” LINK

Eckman, T. 2010. “Some Thoughts on Treating Energy Efficiency as a Resource.” Northwest Power and Conservation Council.

Fulmer, M. and B. Biewald. 1994. “Misconceptions, Mistakes And Misnomers in DSM Cost-Effectiveness Analysis.” ACEEE. LINK

Fuchs, L, L. Skumatz, and J. Ellefsen. 2004. “Non-Energy Benefits (NEBs) from ENERGY STAR: Comprehensive Analysis of Appliance, Outreach, and Homes Programs.” In *Proceedings of the 2004 ACEEE Summer Study on Energy Efficiency in Buildings*. 2.79-2.89. Washington, DC: American Council for an Energy-Efficient Economy.

Hall, N., R. Ridge, G. Peach, S. Khawaja, J. Mapp, B. Smith, R. Morgan, P. Horowitz, G. Edgar, J. Luboff and B. Evans. “Reaching our Energy Efficiency Potential and our Greenhouse Gas Objectives – Are changes to our Policies and Cost Effectiveness Tests Needed?” n.d. Unpublished paper.

Lutzenhiser Associates. 2004 Final Evaluation Report: California Building Performance Contractors Association Comprehensive Whole House Residential Retrofit Program. CPUC-172-02. Portland, OR. Lutzenhiser Associates.

Knight, R.L., B. Knight, L. Lutzenhiser and S. Lutzenhiser. 2006. “Why Comprehensive Residential Energy Efficiency Retrofits Are Undervalued.” Bevilacqua-Knight Inc. & Lutzenhiser Associates. LINK

Jody Myers and Lisa A. Skumatz. 2006. “Evaluating Attribution, Causality, NEBs, and Cost Effectiveness in Multifamily Programs: Enhanced Techniques.” In *Proceedings of the 2006 ACEEE Summer Study on Energy Efficiency in Buildings*. 2.216-2.228. Washington, DC: American Council for an Energy-Efficient Economy.

NAPEE (National Action Plan for Energy Efficiency). 2008. "Understanding Cost-Effectiveness of Energy Efficiency Programs: Best practices, Technical Methods, and Emerging Issues for Policy-Makers." LINK
Neme, C. and M. Kushler. 2010. "Is it Time to Ditch the TRC? – Examining Concerns with Current Practice in Benefit-Cost Analysis." Energy Futures Group and American Council for an Energy-Efficient Economy. LINK

Optimal Energy, Inc. 2011. "Analysis of New York Cost-Effectiveness Screening Methodology and Framework for Energy-Efficiency Programs." LINK

Quinlan, M. 2008. "Cost Effectiveness – Which Test is Best for Screening Conservation Programs?" Connecticut Department of Public Utility Control (Electric).

Regulatory Assistance Project. 2011. *Revenue Regulation and Decoupling: A Guide to Theory and Application*. Montpelier, VT. Regulatory Assistance Project.

Schweitzer, M. and B. Tonn. 2002. "Non-Energy Benefits From the Weatherization Assistance Program: A Summary Of Findings From The Recent Literature." ORNL (Oak Ridge National Lab), prepared for the U.S. DOE. LINK

Skumatz, L., C.A. Dickerson and B. Coates. 2000. "Non-Energy Benefits in the Residential and Non-Residential Sectors: Innovative Measurements and Results for Participant Benefits." In Proceedings of the 2000 ACEEE Summer Study on Energy Efficiency in Buildings. 8.353-8.364. Washington, DC. American Council for an Energy-Efficient Economy.

Titus, E. and J. Michals. "Debating Net Versus Gross Impacts in the Northeast: Policy and Program Perspectives." In *Proceedings of the 2008 ACEEE Summer Study of Energy Efficiency in Buildings*. 5.312-5.323. Washington, DC. American Council for an Energy-Efficient Economy.

Titus, E., M. Nevius and J. Michals. 2004. "How Do We Measure Market Effects? Counting the Ways, and Why It Matters." In *Proceedings of the 2004 ACEEE Summer Study of Energy Efficiency in Buildings*. Washington, DC. American Council for an Energy-Efficient Economy.

Harold Wilhite. 1994. "Market Signals Fall Short as Policy Instruments to Encourage Energy Savings in the Home." In *Proceedings of the 1994 ACEEE Summer Study on Energy Efficiency in Buildings*. 1.193-1.200. Washington, DC. American Council for an Energy-Efficient Economy.

White Paper - Measure it Right: Best Practices in the Selection and Implementation of Cost-Effectiveness Tests

The National Home Performance Council, Inc. (NHPC) encourages the implementation of whole-house retrofits for increased home energy performance and facilitates coordination among federal governmental agencies, utilities, state energy offices, contractors and others to achieve improved whole-house energy performance. NHPC is a registered 501(3)(c) organization. Further information about our firm can be found at www.nhpci.org.

National Home
Performance Council

National Home Performance Council, Inc.
1620 Eye Street NW, Suite 501
Washington, DC 20006
202.463.2005

www.facebook.com/NHPConline

@NHPConline