Building Confidence in Meeting Energy Efficiency Targets – Addressing Free Riders and Additionality

Daniel M. Violette, Navigant Consulting, Inc., Boulder, Colorado, USA
Dimitris Vantzis. Navigant Consulting, Inc., London, UK

Abstract

Building confidence in the estimates of achieved impacts from energy efficiency (EE) programmes is now more important than ever. The European Union Energy Efficiency Directive of December 4, 2012, sets a target of a 20% increase in EE,[1] addressing virtually all aspects of the energy system from supply to consumption. Similarly, regions in North America are also setting aggressive EE targets. For example, the New England states are expecting to meet all demand growth through increased EE efforts.

The treatment of free riders/additionality is one important component in building confidence in EE impacts. The World Energy Congress's (2013) review of energy policies in four European countries highlights this issue: *"As with all public policies, there is a need to address dead-weight/free riders/additionality issues i.e. end users that would have utilised the energy saving measure anyway."*

This paper compares approaches for estimating net savings in Europe and North America. The paper draws upon the U.S. Department of Energy's Uniform Methods Project[2] Section on estimating net savings from EE efforts. For the European perspective, reviews of methods and applications for estimating net savings and attribution employed in Europe and other regions are relied upon. In addition, interviews with evaluation researchers were also conducted to provide additional perspectives and selected evaluation efforts were reviewed.

Introduction

The approaches for estimating overall energy savings and net savings/additional savings have many common elements across Europe and North America. Almost every approach used in Europe is also used in North America; however, there are distinct differences in the philosophy and emphasis in the methods applied. In particular, differences are seen in eight areas:

1) The role of ex ante and ex post approaches for estimating savings
2) Methods for estimating savings from behavioural programs
3) Use of randomized design approaches in evaluation
4) The relative emphasis on free riders versus spillover and market transformation
5) Views on trade-offs between study costs versus the value of information that can be produced by studies focussed on additionality
6) The treatment of self-selection in estimating energy savings has received considerable attention in North America, but is not typically considered in Europe.
7) The focus on statistical approaches designed to achieve certain confidence and precision targets
8) Views on the role of ex ante estimates of net savings and the need for fieldwork to confirm the initial ex ante savings estimates other than validating installation[3]

[1] This target was initially set by the European Union (EU) in 2007. It was reiterated in 2012, with mandatory targets agreed to by member states.
[2] This specifically draws from the section on "Estimating Net Savings – Methods and Practices (Violette and Rathbun, 2014). More detail on the overall U.S. Department of Energy's (DOE's) Uniform Methods Project (UMP) can be found at: http://energy.gov/eere/about-us/initiatives-and-projects/uniform-methods-project-determining-energy-efficiency-progr-0.
[3] Large commercial and industrial (C&I) projects would have savings based on the equipment installed and the

The eight themes listed above are used in this paper to illustrate differences in assumptions and views regarding the value of conducting different types of studies, and in the determination of what constitutes credible evidence between evaluation efforts in Europe and in North America. In addition, the evaluation literature on both continents indicates that different methods are viewed as appropriate for different applications. In the United States, this is demonstrated by the different requirements for addressing net savings and additionality across the different states that regulate electric utilities. In Europe, this is demonstrated by the variety of approaches taken by different countries.

The balance of this paper sets out the estimation approaches, contrasts different approaches with a focus on the eight themes listed above, and presents conclusions on what views (or beliefs) regarding evaluation support different methods, as well as examining directions for future evaluation work that might blend some of the best aspects of efforts in North America and Europe.

Background – Defining the Problem

The goal of investments in energy efficiency (EE) activities is to increase the amount of EE over what would have occurred naturally (i.e., what would have occurred in the absence of the EE activities). This is often termed naturally occurring energy efficiency. Naturally occurring savings is the "baseline," or "the counter-factual scenario," from which EE impacts are measured. As a result, the savings from an EE investment are those that are net of what would have occurred naturally. In North America, energy savings above this baseline are termed net savings. In Europe, these savings are generally termed "additional savings;" however, the term net savings is also used. Net savings can either be estimated directly, or by a calibration of gross savings, where gross and net savings are defined as follows:

- **Gross savings:** Changes in energy consumption that result directly from program-related actions taken by participants of an EE program or activity, regardless of why they participated

- **Net savings:** Changes in energy use that are attributable to a particular EE programme (i.e., they are net or additional to the baseline that represents what would have happened in the absence of the EE program)

Estimates of net savings are often based on first estimating gross savings and making adjustments to gross savings to produce a net savings value. Equation 1 presents a standard approach used in North America, and Equations 2 and 3 are drawn from recent evaluation work in Europe.

Net Savings = Gross Savings – FR + SO + ME (Equation 1)

Where

<u>FR</u> = Free ridership is the programme savings attributable to free riders. Free riders are programme participants who would have implemented programme measures or practices offered by the EE activity even if the EE activity had not been offered.[4]

<u>SO</u> = Spillover savings refers to additional reductions in energy use that are due to EE programme influences that go beyond those directly associated with the EE

characteristics of the equipment replaced based on the specifics of the site.

[4] There are a number of different ways free ridership can influence savings. These can be found in Violette and Rathbun (2014).

activity[5] participation by the end user (i.e., additional energy savings actions not directly part of the programme). As a result, these savings may not be recorded in the EE activity tracking and not credited to the programme.
ME = Market effects are those additional savings not already captured by SO.[6]

Two sets of equations setting out net savings or additionality come from recent work on evaluation in Europe. Bundgaard et al. (2013) set out net savings as:

Net impact = Reported savings * Technical accuracy in the calculation of savings * Additionality * Rebound * Spillover (Equation 2)

Where
- *Technical accuracy* in the calculation of the reported savings refers to the over- (or under-) estimation of the savings due to calculation errors or improper/incorrect use of assumptions.
- *Additionality* occurs if the measure or project would not have been implemented or accelerated without the obligated party's involvement. *Additionality* expresses the likelihood that the energy savings would not have been realised without the obligated party's involvement.
- *Rebound effect* occurs when participants replace the savings achieved with a new or increased consumption of energy. Neither spillover nor rebound effects were quantified in the 2012 evaluation.
- *Spillover* is defined as the "positive co-benefits of energy efficiency programs and measures to promote energy savings."

Vreuls (2012, p.8) sets out net savings using the formula:

*Net Savings = total gross annual savings * f(DC) * f(MP) * f(FR) * (RE)s* (Equation 3)

Where:
- f *(DC)* is double counting;
- *f (MP)* is the multiplier effect;
- *f (FR)* is the free-rider effect; and,
- *f (RE)* is the rebound effect.

All of these net savings equations have common elements; however, they differ in terms of the view of net savings that is taken and the techniques that are used to estimate elements of these net savings relationships. The North American view is that double counting and rebound effects should be addressed in gross savings and are not an adjustment needed to get to net savings. This is rather a small point; however, it is very important to know what the initial reported total or gross savings estimates represent to insure that any adjustments made to a gross savings estimate to arrive at net savings are appropriate.

[5] Spillover can occur both among participants that take actions beyond those offered by the EE programme, but due to the exposure to EE that the programme provided, and to non-participants that become aware of the programme but take actions outside of the boundaries of the programme. See Violette and Rathbun (2012).
[6] Market effects refer to "a change in the structure of a market or the behavior of participants in a market that is reflective of an increase in the adoption of energy efficiency products, services, or practices and is causally related to market" intervention(s)" (Eto et al., 1996). For example, programmes can influence design professionals, vendors, and the market (through product availability, practices, and prices), as well as influence product or practice acceptance and customer expectations.

Another application where additionality is important in is the Clean Development Mechanism (CDM) for achieving reductions in carbon dioxide emissions. For CDM projects, there is a need to demonstrate additionality (i.e., demonstrate that the project would not be implemented without the CDM). Schneider (2007) states that “the demonstration of additionality has been discussed controversially since the establishment of the CDM” and that “the fundamental challenge is that the question as to whether a project would also be implemented without the CDM is hypothetical and counter-factual – it can never be proven with absolute certainty.” Schneider’s analysis of 93 registered projects revealed deficiencies in the estimates of additionality over the past three years, and that “these findings suggest that there are serious problems in the way in which additionality has been assessed.” While the CDM projects are not generally implemented by energy suppliers or utilities as is common for EE obligations in Europe, it highlights the importance of developing appropriate methods and applications.

Methods for Estimating Net Savings

Two recent reports address evaluation methods in Europe and compare these methods to those used in the U.S. and in other regions. A report by Ecofys for the United Kingdom (UK) Department of Energy and Climate Change was “commissioned to explore the needs and requirements for a robust approach to measurement, verification and additionality (M&V and additionality) of electricity demand reduction projects in the context of providing financial incentives for electricity efficiency” (De Lovinfosse et al., 2012). A regional review of best practices prepared by the Regulatory Assistance Project (RAP) as part of the a Global Best Practices series[7] addressed approaches to estimating net savings in Europe and the U.S., as well as in China and India (Slote, et al., 2014). In addition, there have been several reviews of evaluation practices in the U.S. including SeeAction (2012) and DOE’s UMP, which produced a chapter on estimating net savings focused on methods and practices (Violette and Rathbun, 2014).

Approaches to Net Savings – United States

A number of approaches for estimating net savings are outlined in Violette and Rathbun (2014). The authors focused on commonly used methods as well as methods that were currently receiving attention in the evaluation community. An overview of the approaches below belies many nuances in their application, but it is meant to describe the primary concepts used in each approach. Examples of each approach are provided in Violette and Rathbun (2014). These approaches include the following:

- **<u>Randomized control trials (RCTs) and quasi-experimental designs</u>** – These are experimental design approaches that use control groups to help establish the baseline against which EE impacts are calculated. RCTs are viewed as the most accurate approach, as they randomize the energy users in the participant group and the control group. Behavioural programmes can allow for random design as part of the initial programme rollout. The rapid expansion of behavioural programmes and the increased amount of savings being credited to these programs have greatly increased the application of RCTs in the U.S., with over 30 evaluations of this type completed in the past few years. When RCTs are not possible, a quasi-experimental design is viewed as the next best alternative. These designs usually develop a comparison group (i.e., a baseline to approximate the counter-factual) using non-participants, and/or participants but in a pre-EE time period; however, the newest approaches involve a number of different algorithms for matching a participating customer to specific non-

[7] The Global Power Best Practice Series is funded by the Climate Works Foundation (CWF) and RAP.

participating customers. The estimation methods include regression equations, and Difference-in-Difference (DiD) methods. These approaches are increasingly being used in the U.S. to evaluate behavioural programmes, information programmes, and pricing programmes designed to increase efficiency. Generally, these programmes have large numbers of participants and are typically applied to residential/small commercial sector programmes. In general, RCT approaches have not been employed to a great degree in Europe, as behavioural programmes are generally not assigned significant amounts of energy savings; however, examples of billing analyses and quasi-experimental designs are found in Europe.[8]

- **Survey-based approaches** – Survey-based approaches can be a cost-effective, transparent, and flexible method for estimating net savings with a focus on free riders and spillover. They have become one of the most often-used methods in EE net savings estimation. A great deal of effort has been put into good survey design, and understanding the strengths and weaknesses of these methods. Surveys may target up to three types of respondents: (1) programme participants, (2) programme non-participants, and (3) market actors.[9] Survey-based approaches are used in evaluations that start with gross estimates, and then adjust for net savings factors (i.e., free riders, spillover, and market transformation). Survey methods for estimating net savings are not widely used in Europe, although one example can be found in Bundgaard et al. (2013).

- **Common practice baseline approaches** – The common practice baseline approach[10] is receiving attention as a method for estimating net savings in the U.S. In this case, the baseline used as the counter-factual is comprised of estimates of what a typical energy user would have done at the time of the project implementation. This common practice approach (also referred to as a standard practice baseline approach) is under development in several jurisdictions in the U.S. and will certainly evolve as more experience is gained developing common practice baselines in different settings (i.e., for different commercial sectors and for different EE technologies and practices). In general, this method is based on using available information to develop an ex ante estimate of net savings, with limited adjustments based on ex post data and analysis. This approach has certain appealing qualities; however, the trade-offs required need to be clarified, both in terms of potential biases and the real costs associated with applying this approach. This method is similar to the Business-As-Usual (BAU) baseline approach, which is one of the most commonly applied approaches in Europe. The most common concern about this approach raised in peer review comments to the DOE UMP effort (Violette and Rathbun, 2014) involved the potential for self-selection bias. An EE programme that allows energy users to select themselves into the programme may attract those energy users who would have taken the high-efficiency actions anyway. If an EE programme attracted only energy users who were predisposed to install the high-efficiency equipment promoted by the program, then the use of a common practice or BAU approach could overestimate energy savings by not fully accounting for all free ridership. While self-selection has been viewed as an important issue in the U.S., it does not seem to have received much attention in evaluation efforts in Europe.

8 Two examples of statistical evaluations of programmes in Europe can be found in Bundgaard et al. (2013) and Nauleau, M. (2014).

9 Delphi panels are also surveys of a panel of experts, but are presented as a separate method from the category of survey methods as described here.

10 The Common Practice Baseline section gave rise to a number of comments. Some reviewers of the Net Savings chapter of the DOE UMP guidelines (Violette and Rathbun, 2012) did not see this method as parallel to the other methods presented in this chapter, as it focuses on ex ante values of the mean of market behaviour and does not look at ex post information on actions or programme participants. In this context, this approach was viewed as more of an ex ante deemed net savings approach. (See Section 3.7 on deemed net-to-gross [NTG] values.) After considering these comments, the Common Practice Baseline approach was viewed as warranting a separate section due, in part, to the recent attention given this approach to net savings.

- **Market-wide sales data analyses** – A market-wide sales data method can capture the total net effect of the programme, including both free ridership and participant and non-participant "like" spillover. The most common approach is a cross-sectional comparison area method in which post-programme data are compared with data from a non-programme comparison area (or multiple comparison areas) for the same point in time. Thus, evaluators can make a comparison between the change in energy use in the programme area from the pre-programme period to the post-programme period *and* the change in the non-programme area over the same period.

- **Top-down evaluations (or macroeconomic models)** – This method uses energy consumption across a large number of cross-sectional units at the national or regional level. In the U.S., this means that the units of observations have been states or large utilities that have reported energy use data to the U.S. Energy Information Administration. An econometric model regresses the total energy consumption in each of these cross-sectional areas against an indicator of energy efficiency effort (e.g., the amount of money spent on energy efficiency programmes). This approach produced estimates at reasonable levels of confidence and significance at the macro level, but attempts to estimate net savings for a single energy supplier or utility has been problematic. There are those, however, that believe additional effort is warranted in trying to get estimates for sub-regions (i.e., for specific utilities or within a state rather than estimating average values across a large number of states or utilities). This application to small areas requires the disaggregation of energy use within small geographic units as well as the intensity of the energy efficiency effort as the causal variable. This level of disaggregation has been difficult to develop in the U.S. at the state level and utility-specific level, where targets for energy savings are often established and need verification at that level of disaggregation.[11] Top-down methods in Europe generally refer to a different type of analysis than the development of a macro-econometric model that is viewed as the top-down method in the U.S. Rather than a regression-based model, top-down efforts in Europe have often involved structural models or input-output models (Giraudet, 2011).

- **Structured expert judgment approaches** – Structured expert judgment approaches involve assembling a panel of experts who have a good working knowledge of the technology, infrastructure systems, markets, and political environments. This approach is one alternative for addressing market effects in different end-use markets. These experts are asked to estimate baseline market share for a measure or behaviour. In some cases, they are also asked to forecast market share with and without the programme in place. Structured expert judgment processes use a variety of specific techniques (e.g., Delphi analyses) to ensure that the panel of experts specify and take into account key known facts about the programme, the technologies supported, and the development of other influences over time.

- **Deemed or stipulated net savings metrics (NTG ratios)** – A NTG ratio is the ratio of net savings to gross savings. A ratio of .75 would mean that the reported gross energy savings is reduced by 25% to account for net savings factors such as free riders. This approach is used in both the U.S. and in Europe, and is given a little broader discussion here. Deemed or stipulated NTG ratios are predetermined values and do not rely on an ex ante calculation-based approach. Deemed values are often based on previous NTG research that was conducted using at least one of the other methods described in this chapter. NTG ratios may be stipulated when the

[11] The discussion of top-down econometric models and their application to different geographic regions is complex and the reader is referred to Violette and Rathbun (2014) and the additional articles cited in therein.

expense of conducting net savings study or ex ante NTG ratio analyses cannot be justified, or when the uncertainty of the potential results is too great to warrant a study. A recent review of 42 jurisdictions in the United States and Canada (which represented nearly all jurisdictions with ratepayer-funded EE programs) found that 14% use a deemed approach to NTG for commercial and industrial (C&I) programs compared to 50% of the jurisdictions which use an active research approach to developing estimates of net savings factors (Navigant 2013).[12] Typically, evaluators that use the active research approach may not conduct a net savings study every year. For example, it is not unusual in a multiyear portfolio cycle to estimate an NTG ratio for an initial year (or possibly every other year), with deemed values used in the subsequent or intervening years.[13]

- **Historical tracing (or case study) method** – This method involves reconstructing the events (such as the launch of a product or the passage of legislation) that led to the outcome of interest. An example of this is developing a "weight of evidence" conclusion about the specific influence a programme had on the outcome. The historical tracing method traces chronologically a series of interrelated events either going forward from the research point of interest to downstream outcomes, or working backward from an outcome along a path that is expected to lead to precursor events. If all likely paths are followed, forward tracing can capture a relatively comprehensive view of project or programme effects. Because the path leads from a programme event, the connection to the event is assured. Backward tracing usually focuses on a single outcome of importance and follows the trail back through developments that seem to have been critical to reaching the identified outcome. These developments may or may not link back to the research programme of interest (Ruegg and Jordan, 2007). Weiss (1997) suggests historical tracing is similar to theory-driven evaluation and can be viewed as an alternative to classical experimental design. This approach suggests that if the predicted steps between an activity and an outcome can be confirmed in implementation, this matching of the theory to the observed outcomes will lend a strong argument for causality. In other words, if the evaluation can show a series of micro-steps that lead from inputs to outcomes, causal attribution, for all practical purposes, is supported by this approach. The broader literature on evaluation has many applications of historical tracing, but in evaluation energy efficiency projects it has seen limited use and, then, in a supporting role to assess the reasonableness of statistical estimates of net savings. However, there are those that believe historical tracing should play a larger role in energy efficiency evaluation. Historical tracing has some parallels to methods used to assess additionality for CDM projects where the method considers several different points of view that lead to the project being additional including barrier analysis, investment analysis, and a comparison to common practice. (UN Framework CDM, 2012; Öko-Institut, 2007.)

The approaches set out above are representative of current methods used in North America but not exhaustive. For example, Violette and Rathbun (2014) devote an appendix to a method that is relatively new and has only been applied in one sector of the energy efficiency market – upstream residential lighting programmes. This approach looks at the price of high-efficiency lighting products and then estimates what the prices would have been if the incentives programme had not

12 Approximately one third of the jurisdictions did not adjust gross savings for either free ridership or spillover; however, many of those states conducted some NTG research to inform future program design. This reflects policy decisions in each state. Several states that did not adjust gross savings for net savings factors at the time of this study have changed or are contemplating changing to approaches that do estimate net savings. Pennsylvania and Maryland fall into this category.

13 Another issue raised by a reviewer was that the use of deemed NTG values can remove the incentive for the program administrator to reduce free ridership and maximize spillover and market effects to yield greater net savings values.

been offered. An econometric model is used to estimate price elasticity based on the range of current market prices, and this price elasticity is used to estimate the incremental sales of lighting products due to the incentives programme.

Approaches to Net Savings – Reviews of Applications in Europe

The basic concepts are similar between the U.S. and Europe. Bertoldi and Rezessy (2009) state that "to determine the energy savings resulting from an energy efficiency activity, the eventual energy consumption has to be compared to a baseline – a counterfactual reference scenario. The choice of the reference scenario – in terms of reference consumption and conditions – raises some challenges, and the practicality and cost-effectiveness of a baseline methodology." The authors indicate that the most common baselines used in the existing Energy Efficiency Obligation EEO schemes are the following:

1. "Sales average" performance

2. Performance of the most commonly used appliance on the market ("average-on-the-market" for appliances and equipment)

The sales baseline is on new equipment, while the definition of the average-on-the-market seems to be a building stock number resulting in a lower baseline and a higher energy savings value.[14] These two baselines are commonly used for appliances and equipment that have a large number of installations, and the net energy savings comes from an ex ante estimate that takes into account both the performance of the equipment and any adjustments suggested by the baseline. However, no documents were found in this limited literature search that documented how net savings factors such as free riders influence this engineering-based calculation of ex ante savings. As a result, this paper cannot shed any light on how this is done.

Most of the factors that would adjust gross savings to a net savings value in this review of methods used in Europe seem to be accounted for in the ex ante engineering estimates of measure savings and for large project on-site savings estimates. As a result, it is useful to look at the three baseline methods used in Italy. These three methods are presented in Bertoldi and Rezessy (2009, p.30):

1) The deemed savings approach does not require in-field measurement. Deemed savings apply to technologies for which energy savings are well known and do not exceed 25 toe[15] per year; default factors for free riding, delivery mechanism, and persistence have been introduced. Examples of measures that are certified using this approach include compact fluorescent lamps (CFLs), m^2 insulated wall, small photovoltaic (PV) applications, and high-efficiency boilers.

2) The engineering approach implies some on-field measurement and applies to measures for which energy savings are known; however, they may differ, depending on a number of restricted factors (e.g., number of working hours). This approach applies to measures that yield up to 50 toe per year (for energy service companies [ESCOs] and small distributors) and 100

[14] Slote et al. (2014) indicate that the baselines used for determining energy savings are not uniform among Member States. This can have significant effects on reported savings. For example, a Member State may use as a baseline the minimum efficiency level currently available for an appliance in the marketplace, or the average efficiency level of models currently sold, and compares this with the savings from the most efficient available model. Another Member State may use as a baseline the efficiency level of older models being replaced.

[15] This refers to Tonnes of Oil Equivalent (toe). The International Energy Agency (IEA)/Organisation for Economic Cooperation and Development (OECD) defines one *toe* to be equal to 41.868 gigajoules or 11.63 megawatt-hours.

toe per year for distributors, respectively.

3) A metered baseline method applies to measures for which energy savings need to be addressed on a case-by-case basis. It entails direct measurement of energy use, and preapproval of proposed baselines and methodologies. This approach applies to measures that yield up to 100 toe per year (for ESCOs and small distributors) and 200 toe per year for large distributors in savings, respectively.

A summary of methods applied across Italy, Great Britain, and France are provided in EuroWhiteCert (2007) and summarized in Table 1. A similar tabulation is presented in Bertoldi and Rezessy (2009, p.31).

Table 1. Overview of Measurement and Verification Approach in White Certificate Systems in France, Great Britain, and Italy

Country	Measurement and Verification System in Place
Italy	AEEG, the Italian Regulatory Authority for Electricity and Gas, uses three evaluation approaches: • Default value: energy saving is defined ex ante • Engineering approach: on-field measurement • Energy monitoring plan
Great Britain	• The regulator OFGEM assesses and approves all measures suppliers take. • DEFRA (Environment Ministry) developed a 'Target-setting Model" for ex ante determining the energy savings attributed to different measures.
France	• ADEME (French Agency for Environment and Energy Management) and ATEE (Association Technique Energie Environment) are in charge of setting methodologies for calculation of the achieved savings. • Savings are validated by the French High Council for Energy.

Source: EuroWhiteCert (2007)

Ecofys (2014) developed a recommendation for robust measurement and verification (M&V) and attribution for a UK incentives programme drawing on information gained through Europe's EEO system. In this recommended UK system, an entity (usually energy retailers or distributors) is assigned an obligation to save energy in eligible end-use customers. If the savings are not met, the obligated entity may pay a penalty. The EEO systems currently in place within the European Union show considerable variation with regard to the obligated actors, the eligible customers, and the eligible energy conservation measures. There is also considerable diversity in how the targets are set and how companies choose to achieve those targets.

The EEO schemes have mainly (but not only) targeted energy efficiency measures that are highly replicable on a large number of projects. Additionality is a concern for the cost efficiency of these policies; therefore, it is intrinsically dealt with when choosing the portfolio of energy efficiency measures and technologies eligible for the scheme upfront. The solution adopted in most EEO schemes is that additionality is defined ex ante at the programme level and reviewed on a regular basis to account for technology, market, and policy changes.

Under the French system, additionality is assessed ex ante for over 100 standardized measures. The decision on additionality of standardised measures is based on technology and market analysis of the savings expected for these measures. Project developers can also suggest new measures and guidelines are available for the so-called non-standardised measures. The lists of eligible technologies by country have a lot of variability among countries. As EEO schemes rely to a large extent on *standard* measures, the selection of measures is typically done ex ante by the Ministry and/or its advisory bodies.

An effective and robust approach to M&V and additionality of electricity savings is always a trade-off between the costs of the approach and the certainty of the savings achieved (Ecofys, 2014). This effort was focused on the UK's need to establish the additionality requirements for a financial incentives programme to promote energy efficiency.

The solution proposed by Ecofys is a deemed ex ante approach where there is a "regular update of the measures supported (based on market, technology and policy analysis)." This is due to the expected higher costs of ex post approaches than greater transparency in the financial transactions. It is also due to the acceptance of some likely non-additional savings (i.e., trying to strive for 100% of additionality is likely to be unreasonable). This view is also common in the U.S., where some degree of free ridership is expected in almost every program. Still, the view in the U.S. is that the value of ex post studies of energy savings – both engineering gross savings and net savings – is worth the cost as in-field gross savings can vary dramatically from those derived from desk engineering.

In summary, the baselines used to assess gross and net energy savings in Europe are largely based on engineering methods and the net savings factors that would reduce (or increase in the case of spillover) energy savings that are also determined ex ante. There is evidence that information has been used to adjust the gross savings numbers based on market circumstances. For example, the UK reduced certain residential measures by 50% due to what was being seen in the market. Now, they have increased these back to nearly 100% due to changes in the equipment included in the program. Italy also reduced gross savings from CFLs by 42%. So, baselines are clearly given consideration, but the documentation of the processes was not reviewed for this paper.

SUMMARY

The eight issues presented in the introduction are used to present a summary of the approaches and views used in net saving research, and how there seem to be subtle, but important, differences in the areas of focus between work in North America and Europe.16 Each of the eight issues is discussed below:

1) **The role of ex ante and ex post approaches for estimating savings**.
The work in Europe, in general, demonstrates a greater confidence in the ability of engineers to develop ex ante estimates of savings across a wide range of measures than is generally found among North American regulators and energy efficiency programme administrators. Generally, deemed savings is used as a programme planning tool, but there still is an ex post study to confirm and validate the deemed savings using in-field data and information. This stems, in part, from the development of EE technical resource manuals in North America and issues identified when the engineering algorithms have been compared to a sample of in-field estimated savings.[17] As a result, the view in the North America, for the most part, is that net savings factors need to be estimated ex post. This does not mean that a net savings study is conducted for every programme every year, but

[16] The authors had wider access to papers in the U.S.; more time devoted to expanded search for papers in Europe might have resulted in a wider range of approaches to additionality being found, rather than relying on survey papers of approaches.

[17] Bundgaard (2013) reviews a set of studies in Denmark using in-field survey approaches and statistical methods. These ex post studies generally produced estimates of net savings in the range of 40% to 60% of the ex ante engineering estimates. However, care is needed in drawing implications from these results for other regions and engineering estimates of energy savings. Engineering estimation methods to produce ex ante estimates may incorporate different elements that can address components of free ridership and other factors that influence net savings. Still, the magnitude of the adjustment to the ex ante engineering savings needed to produce net savings estimates in these studies, could be seen as implying that additional in-field verification of ex ante engineering estimates may be warranted. (See also Togeby, M., et al., 2012a and 2012b for additional discussions of the evaluations in Denmark.)

the general rule is to conduct these studies every two to three years, with possibly an interim confirmatory validation study on a small sample to see if the net savings values from past studies are still appropriate. New developments in sample design based on the characteristics of energy savings and energy use data have lowered the costs of in-field studies to a point where most regulators and programme administrators believe that the value of information from these ex post studies is greater than the costs incurred. This is not universally agreed upon in North America; however, in the authors' opinion, it depicts the current evaluation environment.

2) **Methods for estimating savings from behavioural programmes.**
Behavioural programs in Europe are generally not viewed as providing direct savings. In North America, recent work has shown substantial savings from programs that provide information on an energy user's bill, ways to reduce their bill, and comparative information regarding how their building or dwelling fares compared to other matched units' energy bills. The somewhat surprising magnitude of the savings from these behavioural programs in North America has resulted in numerous evaluations of behavioural programs designed to address additionality and produce accurate estimates of net savings using statistical approaches.

3) **Use of randomized design approaches in evaluation.**
The use of RCTs has expanded greatly in recent years as some energy efficiency programmes have been made opt-out rather than opt-in programmes. Usually, these are innovative rates or information programmes that are applied to a large population of energy users. In these opt-out programs, customers are in the programme unless they make a proactive choice not to participate and inform the programme administrator. The recognition of the advantages of RCTs has made more programme designers and administers willing to build in the random assignment of a control group as part of programme rollout.

4) **The relative emphasis on free riders versus spillover and market transformation.**
In assessing net savings factors, both North America and Europe emphasize free riders over the other factors that comprise net savings (i.e., spillover and market transformation). However, a growing number of jurisdictions in the U.S. are requesting information on these two factors, even if it is directional and more of a magnitude estimate than a precise point estimate. As a result, the evaluation community is developing approaches using panel survey data and trade ally surveys to address these issues.

5) **View on trade-offs between study costs versus the value of information that can be produced by studies focused on additionality.**
Ecofys (2014) and other surveys of evaluation, measurement, and verification approaches in Europe emphasize the cost versus value trade-offs in performing ex post energy savings studies. This is also reflected in the work in North America, where Violette and Rathbun (2014) state that "some methods that provide more credible results are costlier. This cost may be justified for programme components that are important to the portfolio, but not for all components." In addition, the systematic assessment of the value of information gained by net savings estimation approaches as compared to the cost of the research is needed to better balance requests to meet confidence and precision levels for net savings estimates. Given the similar stated objectives, Europe spends considerably less on evaluation as a percentage of energy efficiency expenditures than is common in North America. The costs of ex post studies balanced against the value of the information produced is a policy and value assessment that should continue to receive consideration in both North America and Europe.

6) **The treatment of self-selection in estimating energy savings has received considerable**

attention in North America, but is not typically considered in Europe.
An issue that has been given considerable attention in North America is self-selection. Self-selection implies that when customers are given a choice to participate in a programme or not, those that do choose to participate may be systematically different than those that choose not to participate. In general, the assumption is that those that choose to participate in an energy efficiency programme are those same energy users that likely would have selected higher efficiency equipment even if the programme had not been offered. This is important in that, if true, it would result in higher free ridership and lower net savings estimates than one would get using a market or business-as-usual baseline. The emphasis that evaluators have given to self-selection in North America has made them wary of the common practice baseline approach that has been proposed in several jurisdictions. There is a belief that the market baseline may not represent the typical participant that selects themselves into an energy efficiency programme. As a result, comments on the inclusion of the common practice baseline approach to net savings received on DOE's Uniform Methods Project ranged from statements that it should not be included as it was not an evaluation method, to maybe this might be a practical approach. The wide use of market-based and sales-based baselines in Europe makes this an issue of note. However, the only way to address this question is to conduct some research on the potential importance of self-selection. A reasonable research agenda should be able to determine if this is a significant factor in programme participation.

7) **North America has a greater focus on statistical approaches to achieve certain confidence and precision targets, with certain jurisdictions specifying the levels of confidence and precision to be obtained in evaluation efforts.**
Evaluation research in North America often has targets in terms of confidence and precision around various components of net savings, and results of surveys. This leads to a focus on efficient sampling strategies and analysis methods. However, there is a growing debate on what these confidence and precision targets should be for EE programme evaluation. Many evaluators believe the regulators and institutions have set targets so high that the research methods may be distorted in an attempt to meet these specific targets. Conventional thinking about confidence and precision targets for estimates of energy programme savings are being reassessed. The traditional 90% confidence and 10% precision targets are being adjusted based on factors linked to the relative contribution of energy programme, the value of information in making decisions about investments in EE programmes, and comparisons to the uncertainties in the production of energy from supply-side resources. This is a substantive discussion in North America. On the other hand, there is less discussion of confidence and precision in Europe, which may be due to the greater use of ex ante deemed savings approaches and less emphasis on ex post statistical approaches.

8) **Different philosophical views on the estimation of energy savings.**
In summary, the literature reviewed indicates that there are different philosophies between North America and Europe towards EE evaluation in general, and in the methods to be used for assessing additionality and net savings in particular. Europe's consideration of these matters has led them to spend less on evaluation, and perform fewer in-field ex post monitoring, metering, and verification studies. (Note: These are not the M&V efforts made on-site at the time of installation to estimate gross savings, but are efforts taken some time later to see if the expected savings are actually being realized). In addition, there have been fewer studies that address energy user characteristics that would contribute to ex post net savings being different than the ex ante estimated savings. In North America, ex post studies have generally produced lower net savings estimates than were anticipated by ex ante estimates, although progress has been made to reduce this difference.

In summary, the goal for evaluation is to provide the information needed for making good decisions regarding spending the public's monies on investments in energy efficiency. Evaluators and programme designers recognize the same issues, yet North America and Europe have taken different paths. Will one path converge towards the other, or will they continue on in divergent directions? Time and the results of what should be a structured research agenda will help guide future courses of action. Are the energy efficiency measures being selected for inclusion in programmes in Europe of a type that make ex ante deemed savings and market baselines (or business-as-usual baselines) appropriate for estimating net savings, or might it be important to confirm these estimates using ex post studies as is more common in North America? The answer will depend in part on results from future evaluations, but it also is a policy question that takes into account the cost and value of the information produced.

References

Bertoldi, P. and S. Rezessy. 2009. "Energy saving obligations and tradable white certificates." Report prepared by the Joint Research Centre of the European Commission. http://ec.europa.eu/energy/efficiency/studies/doc/2009_12_jrc_white_certificates.pdf.

Bundgaard S.S., M. Togeby, K. Duhr-Mikkelsen, T. Sommer, V. Hansen-Kjaerbye, and A. Larsen. 2013. "Spending to Save: Evaluation of the Energy Efficiency Obligation in Denmark". *In Proceedings of the ECEEE 2013 Summer Study on Energy Efficiency*. Toulon, France. European Council for an Energy Efficient Economy. http://proceedings.eceee.org/visabstrakt.php?event=3&doc=7-131-13.

De Lovinfosse, I., L. Jansiro, K. Blok, and J. Larkin. 2012. "Measurement, Verification and Additionality of Electricity Demand Reductions." Prepared by Ecofys. *https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/66216/Electricity_demand_reduction-_measurement__verification_and_additionality.pdf.*

Eto, J., et al. 1996. *A Scoping Study on Energy-efficiency Market Transformation by California Utility DSM Programs*. Lawrence Berkeley National Laboratory. http://emp.lbl.gov/sites/all/files/lbnl%20-%2039058.pdf.

European Commission. 2012. *Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency*. http://www.buildup.eu/publications/32236.

EuroWhiteCert. 2007. *Package 4.1: Supply side: measurement and verification of energy efficiency projects.* Report presented to Intelligent Energy Executive Agency by Armines and ISR-UC on behalf of the EWC team.

Giradut, L. 2011. Exploring the Potential for Energy Conservation in French Households through Hybrid Modeling. CIRED Working Papers No 26-2011. http://www.centre-cired.fr/IMG/pdf/CIREDWP-201126.pdf.

Navigant. 2013. *Custom Free Ridership and Participant Spillover Jurisdictional Review*. Prepared for the Sub-Committee of the Ontario Technical Evaluation Committee, May. www.ontarioenergyboard.ca/documents/TEC/Evaluation%20Studies%20and%20Other%20Reports/Ontario%20NTG%20Jurisdictional%20Review%20-%20Final%20Report.pdf.

Ruegg, R. and G. Jordan. 2007. *Overview of Evaluation Methods for R&D Programs*. Prepared for U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy.

Schneider L. 2007. “Is the CDM fulfilling its environmental and sustainable development objectives? An evaluation of the CDM and options for improvement.” *Öko-Institut.*

SeeAction – State and Local Energy Efficiency Action Network. 2012. *Energy Efficiency Programme Impact Evaluation Guide*. Prepared by Steven R. Schiller, Schiller Consulting, Inc., www.seeaction.energy.gov.

.Slote S., M. Sherman, and D. Crossley. 2014. *Energy Efficiency Evaluation, Measurement, and Verification*. Regulatory Assistance Project (RAP) Global Power Best Practice Series. http://www.raponline.org/topic/global-power-best-practice-series.

Thomas, Stefan. 2009. *Measuring and reporting energy savings for the Energy Services Directive – how it can be done – Results from the EMEEES project*. Wuppertal Institute on behalf of the EMEEES Consortium. http://www.evaluate-energy-savings.eu/emeees/en/publications/reports/EMEEES_Final_Report.pdf.

Togeby, M., et al. 2012. "A Danish Case Study: portfolio evaluation and its impact on energy efficiency policy." *Energy Efficiency* 5: 37–49. http://link.springer.com/article/10.1007%2Fs12053-011-9117-7#page-1.

.

UNFCCC - United Nations Framework Convention on Climate Change. 2012. “Approved Baseline and Monitoring Methodologies for Large Scale CDM Project Activities”. http://cdm.unfccc.int/methodologies/PAmethodologies/approved.

UNFCCC – United Nations Framework Convention on Climate Change. 2011. “Tool for the demonstration and assessment of additionality”. http://cdm.unfccc.int/methodologies/PAmethodologies/tools/am-tool-01-v7.0.0.pdf.

Vreuls, Harry. 2012. *A new tool for evaluators: the European standard on energy savings calculations*. International Energy Policies & Programmes Evaluation Conference. Rome, Italy. June 12-14, 2012.

Violette D. and P. Rathbun. 2014. *Estimating Net Energy Saving: Methods and Practices*. Prepared for the The Uniform Methods Project: Methods For Determining Energy Efficiency Savings, U.S. Department of Energy. Project Site with specific document link forthcoming: http://www.energy.gov/eere/downloads/uniform-methods-project-methods-determining-energy-efficiency-savings-specific

Weiss, C. (1997). “Theory-Based Evaluation: Past, Present, and Future” Special Issue: Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods, New Directions in Evaluation, Volume 1997, Issue 76.

World Energy Council. 2008. *Energy Efficiency Policies around the World: Review and Evaluation*. London, UK.