EPATEE

Comparing Estimated versus Measured Energy Savings

EPATEE topical case study illustrated with examples

Table of contents

Summary

This topical case study compares two approaches to evaluate energy savings: detailed engineering estimate and evaluations based on measurement, which could be a direct measurement (often within the industrial sector) or a billing analyses (often within the residential sector). In essence this case study looks at the difference between 'estimated' versus 'actual measured' energy consumption for the baseline and the situation after the energy efficiency actions have been implemented. When reliable measured or metered data of energy consumption can be collected for all participants and that external factors can be controlled, measurement-based approaches will likely provide more reliable results. But in practice, these conditions are far from being commonly met. This topical case study therefore discusses the pros and cons of both approaches and the cautions to take when comparing their results. This topic is a general issue, relevant for all sectors and types of policy measures.

We clarify first the different types of data and methods used in energy savings evaluations and considered in this case study. Among the different evaluation methods distinguished in EPATEE, the ones relevant for this case study are the five bottom-up methods: direct measurement (method 1), billing analysis (method 2), deemed estimate of unitary energy savings (method 3), mixed deemed and ex-post estimate (method 4), and detailed engineering estimates (method 5).

We then review when to use each approach ('estimation-based' and 'measurement-based') according to literature. This neither is a black-and-white situation: methods are often used in combination with each other, very much depending on data availability. Larger programs tend to need more data and more types of data for evaluation purposes; especially when factors as the rebound effect are investigated.

A synthesis concludes the literature review by explaining what main factors can cause differences between estimated and measured energy savings, and how they can be analysed and taken into account. References are provided for readers looking for examples or more details.

Finally, two practical examples illustrate the topic further. For the Better Energy Homes scheme in Ireland, *both* estimation-based (engineering calculations) and measurement-based (billing analysis)

The EPATEE project has received funding from the European Union's Horizon 2020 Research and innovation programme under grant agreement No 746265.

approaches have been used. The engineering calculations are used to monitor and report energy savings on a regular basis, providing a continuous feedback on the results of the scheme. The estimation-based approach makes possible to calculate savings based on the data that can easily be collected from each participant (e.g., type of actions installed, type of buildings). They represent 'theoretical' energy savings as they are based on standard assumptions (e.g., about energy behaviours or performance of the actions). Then billing analysis was used for an ex-post evaluation that aimed at investigating the energy savings actually achieved by the participants, and to what extent energy savings could be attributed to the scheme (using a control group). This required tremendous efforts in collecting billing data. The difference between the two approaches is considered as the 'loss of savings' (or shortfall[1]) due to e.g. a rebound effect or poorly installed equipment. The evaluators acknowledged that the difference in the results from the engineering estimates and the billing analysis could not be disaggregated into the potential influences of the various factors that might affect these results. In practice this means that billing analysis does not always deliver perfect results or perfect understanding of the results, unless data collection was planned early enough. This would indeed require to collect many complementary data[2]. The evaluators also mentioned that despite all their efforts put in data collection and rigorous data analysis, there were still some bias that could create uncertainties in the results from the billing analysis.

The second example is about a renovation programme in Lithuania. Difficulties in collecting metered data were also encountered, restricting the opportunities to form a sample to compare estimated and measured energy savings. While for Ireland the data were collected from the gas network operator, for Lithuania the data were collected from a district heating company. In both cases the estimated energy savings were based on simplified engineering calculations as defined for the respective national Energy Performance Certificates. In both cases, the measured savings were about 36% lower than the estimated savings.

Scope and definition

The EPATEE project deals with the **evaluation of energy savings from policy measures**, with a focus on **bottom-up methods**. These methods start by evaluating the energy savings at the level of the participants to the policy measure (or sample(s) of participants), and then sum up (or extrapolate) the results over the total number of participants to get the total energy savings of the policy measure.

The review of the evaluations gathered in the EPATEE Knowledge Base or analysed in the EPATEE case studies shows that in most cases, energy savings were evaluated either based on methods based on engineering estimates (with various levels of simplification/sophistication) or on methods based on measured or metered data (mostly energy bills, or direct measurements to a lesser extent).

The literature shows that large discrepancies can be found in the results from the two approaches ('**estimation-based**' and '**measurement-based**'), which was also found in the EPATEE case studies where evaluations have compared both approaches (as illustrated by the example n°1 from Ireland).

The objective of this topical case study is thus to **discuss the differences between both approaches, their pros and cons, and how they can be compared**. This topic is a general issue, relevant for all sectors and types of policy measures.

1 This term is used by the authors of the paper about the billing analysis (Scheer et al., 2013).
2 Another EPATEE case study, the 'Warm Front' in the UK, illustrates this topic further (Broc, 2018).

The comparison between 'estimation-based' and 'measurement-based' approaches is here considered **when evaluating gross energy savings** (i.e. from the point of view of the end-users, not taking into account causality or attribution). Differences can also occur when evaluating net energy savings (i.e. from the point of view of the policy maker or implementer, taking into account causality or attribution). The issue of evaluating net energy savings is the subject of another topical case study (see Voswinkel et al., 2018).

This document focuses on the **comparison between bottom-up methods**. But the issue can also arise when using top-down methods or mixed methods.

Insight from the literature

Types of data and methods for a bottom-up evaluation

When evaluating energy savings from a policy measure with bottom-up methods, the following types of data are often needed (Eichhammer 2008):

1. The number of participants that have joined the program
2. The number and types of actions taken by these participants
3. The energy consumption for each action during a certain period of time, compared to the baseline energy consumption[3]. The consumption could be measured within the programme, taken from other sources as 'deemed' (e.g., based on previous studies, laboratory tests, etc.), or calculated from other variables (see for example the calculation methods from the EMEEES project for common types of energy efficiency actions[4]).
4. Normalisation factors[5] correcting e.g. for weather conditions (heating degree days), production volumes, etc.
5. Adjustment factors, such as double counting or free-riders (see EPATEE, 2017 for a list of effects and related definitions as used in EPATEE), to go from gross energy savings (point of view of the end-users) to net energy savings (point of view of the policy measure)[6].
6. Data to calculate indirect effects (negative and co-benefits) of energy savings, such as job creation in certain sectors[7]

The types of data that are of interest for this case study are data types 2 to 4.
Type 1 is a requirement for all bottom-up methods. It is usually collected through the regular monitoring of the policy measure. When the monitoring keeps track of the number of participants

[3] See (EPATEE 2017) about the typology of baselines as used in EPATEE. For more general discussions about the baseline issue, see for example section 3.3 of (SEEAction, 2012).

[4] http://www.evaluate-energy-savings.eu/emeees/en/evaluation_tools/bottom-up.php

[5] Normalizations factors are sometimes called 'adjustments'. They bring energy use in the two time periods (baseline situation and after the introduction of an energy savings action) to the same set of conditions. Conditions commonly affecting energy use are weather, occupancy, plant throughput, and equipment operations required by these conditions. Adjustments can be positive or negative.

[6] About this, see the topical EPATEE case study about evaluating net energy savings (Voswinkel et al. 2018). The concepts of gross and net energy savings, free riders and spill-over is also, amongst other reports, discussed in section 5 of the European Ex-post Evaluation Guidebook for DSM and EE Service Programmes (SRCI et al., 2001),section 5 of (SEEAction, 2012) or Chapter 21 of the protocols of the Uniform Methods Project (Violette and Rathbun, 2017).

[7] Section 7.9 of (SEEAction, 2012) describes non-energy benefits. Input-output analysis (included in method 9 in the EPATEE typology) is for example often used to assess impacts on employment.

and/or actions, it is not a direct source of differences in the results between the different types of evaluation methods. Uncertainties can arise from the use of surveys to estimate the number of participants or actions. But then they occur independently of the way energy savings per participant (or action) are calculated.
Type 5 is related to the evaluation of net energy savings, which is dealt with in another topical case study (see Voswinkel et al. 2018).
Type 6 is related to the evaluation of non-energy impacts, which is also out of scope for this case study.

In EPATEE we have distinguished 10 evaluation methods (see EPATEE, 2017), based on the typologies previously defined in the EMEEES project and the MURE database[8]. Five of these methods are bottom-up methods, thus in the scope of this topical case study. The table below presents these methods, and more specifically their general data requirements related to the number and type of actions (data type 2), determination of energy consumption (data type 3) and the normalization factors (data type 4). Understanding the differences in these data is indeed essential to discuss the comparison between 'estimation-based' and 'measurement-based' methods. Moreover, the table also reminds the type of output for each method, according to the typology of energy savings defined in the Annex V of the EU Energy Efficiency Directive (2012/27/CE).

Table 1. Bottom-up methods distinguished in EPATEE: main input and output data related to evaluation of gross energy savings.

#	short description	Method to collect input data	Output
1	**direct measurement (of energy consumption)**	**Number/type of actions**: monitoring needed if measurements made on samples (then data needed for extrapolating results from the sample to the whole participants) **Energy consumption**: metering or sub-metering, laboratory test (usually done on samples of equipment or systems) **Normalisation factors**: needed if the metered energy end-use can be influenced by external factors and that the measurement period is not long enough to provide average data representative of typical conditions (then complementary monitoring of relevant variables can be needed) ***Example***: energy consumption from laboratory tests used to assess energy savings from policies to implement minimum energy performance standards (see the EPATEE case study on Nordsyn: Dragovic and Broc, 2018)	Metered savings
2	**billing analysis**	**Number/type of actions**: monitoring needed if energy bills collected for samples (then data needed for extrapolating results from the sample to the whole participants) **Energy consumption**: energy bills (or metered data) collected from the participants or their energy utilities **Normalisation factors**: needed if the external factors cannot be controlled by a comparison with a control group or with long enough time series (then complementary monitoring of relevant variables) ***Example***: more details about using this method can found in the EPATEE case studies about Better Energy Homes (see example	Metered savings

[8] EMEEES: http://www.evaluate-energy-savings.eu ; MURE database: http://www.measures-odyssee-mure.eu/

		n°1), Warm Front (Broc, 2018), or the Weatherization Assistance Program (Spyridaki and Broc, 2018).	
3	**deemed estimate**	**Number/type of actions**: monitoring needed **Energy consumption**: review of data sources available (e.g., previous studies, manufacturer data, national statistics) either for a direct estimate or for a simplified engineering calculation **Normalisation factors**: deemed estimates are made for normalised conditions, so no need to apply normalisation factors afterwards ***Example***: schemes using this method can be found in many of the EPATEE case studies, for example, Individual heat metering in multifamily buildings in Croatia (Bukarica et al. 2018) or the subsidy scheme for housing corporations in Amsterdam (Meulen, J. and M. Menkveld 2018).	Deemed savings
4	**mixed deemed and ex-post**	**Number/type of actions**: monitoring needed **Energy consumption**: mix of review of data sources available, and data collected from the participants for data that can vary and influence significantly the energy savings (data collection usually based on paper or online forms, and/or on invoices) **Normalisation factors**: simplified engineering calculations are usually made for normalised conditions ***Example***: schemes using this method can be found in several of the EPATEE case studies, for example, Better Energy Homes (see example n°1), or the Italian white certificates scheme (Di Santo and Biele 2017).	Deemed savings
5	**detailed engineering estimates**	**Number/type of actions**: monitoring needed if the detailed estimates are done on a sample (then data needed for extrapolating results from the sample to the whole participants) **Energy consumption**: data collected from the participants (usually through an energy audit or alike) **Normalisation factors**: detailed engineering calculations or modelling are usually made for normalised conditions ***Example***: two typical examples for this method are the use of energy audits or modelling software (e.g., to assess energy consumption of buildings). See for example the EPATEE case studies about Energy renovation of public sector buildings in Croatia (Marić et al. 2017), Energy Efficiency Agreement for Industries in Finland (Gynther and Suomi 2017), Renovation programmes in Lithuania (see example n°2).	Scaled savings

Engineering estimates versus measured data; what is it?

This topical case study looks at the difference between evaluation methods based on engineering estimates versus measured data, which can be read as 'theoretically calculated savings' versus 'actual measured savings'. Choosing what evaluation method to use is based on several factors. Some of these may include (TeckMarket Works et al. 2004):

- Type of policy instrument
- Data collection costs and available budget

- Size of the policy or programme
- Programme or policy evaluation history (verify/compare/triangulate with another method, questions raised from past evaluations, etc.)
- Demonstrated effectiveness of methods when applied to other similar policies or programmes
- Whether the evaluation is of a single policy or a consolidated evaluation of several policies

The choice of the evaluation method also needs to be consistent with the evaluation objectives, and thereby with the policy objectives. For example, when evaluating a policy that promotes technical energy performance improvements, evaluators might choose to use an estimation-based approach, as the measurement-based approach will also capture changes in energy consumption due to other factors such as energy behaviours. At the opposite, evaluators might choose to use a measurement-based approach for the same reason when evaluating a policy that aims at reducing energy consumption or GHG emissions.

For methods based on **engineering estimates**, (Vreuls et al. 2009) distinguishes two sub-methods:

1. **Simple engineering calculations** (deemed estimates, or mixed deemed and ex-post in the EPATEE typology): calculations *without* on-site inspection is a common method for obtaining data for estimating energy savings. Data may be estimated using engineering principles, without using on-site data, but with assumptions based on equipment specifications, performance characteristics, operation profiles of devices installed, etc.
2. **Enhanced engineering calculations** (or modelling) (detailed engineering estimates in the EPATEE typology): energy consumption data may be calculated on the basis of information obtained by an external expert during an audit of one or several targeted sites, or through information collected from the participants or installers (e.g., using online forms). On this basis, more sophisticated algorithms/simulation models could be developed and be applied to a larger population of sites (e.g. buildings, facilities, vehicles). Measurement can be used to complement and calibrate the algorithms/simulation models.

For methods based on **measurements**, the measured energy consumption could come from several sources:

1. **Measuring or metering the end-use load data** (direct measurement in the EPATEE typology): state-of-the-art guidelines about these measurements can be found in (Mort 2017)[9]. Measured data can be directly the energy consumption, e.g. the electricity consumption of a device during operating hours. In that case, it is equivalent to method 1 in the EPATEE typology. Measured data can also be done on variables used to calculate energy consumption, e.g., power or operating hours. In that case, it provides data to method 4, where the measured data will be the ex-post data that might be combined with other data sources (ex-ante estimates). Measuring can be done continuously or during a short-term. It can be relatively expensive and mostly applicable for a small

[9] "*Metering is defined as the use of instrumentation to measure and record physical parameters. In the context of energy-efficiency evaluations, the purpose of metering is to accurately collect the data required to estimate the savings attributable to the implementation of energy efficiency measures (EEMs). Estimated energy savings are calculated as the difference between the energy use during the baseline period and the energy use during the post-installation period of the EEM*". Mort (2017) describes metering methods for several types of measurements, discussing skill-level requirements, measurement considerations (e.g., timing, length, periodicity), metering equipment types and their respective measurement accuracies, data handling procedures.

number of projects or for samples of projects, unless measurements are already done for other purposes[10]. What will be measured is determined by the user: this could be one device, a group of devices, one industrial plant or a complete building of facility.

2. **Making use of smart meters** (intermediate between direct measurement and billing analysis, or specific type of billing analysis in the EPATEE typology), providing e.g. the energy consumption each hour. Smart meters can be installed to measure energy consumption of a group of devices, part of or a whole building or site, during short time intervals, whole year round. Their data are used for energy billing. But they can also provide data with a shorter timespan or load profiles.
3. **Collecting energy bills** from distribution or supply companies or from the participants (method 2); showing the yearly consumption of what is connected behind the meter. The timespan can be shorter than one year, depending on the frequency of meter reads. In the evaluations analysed for the EPATEE case studies, energy bills were the most frequent data source when a measurement-based method was used (as shown in the two examples of this case study).

It should be noted that smart meters and energy bills are mostly available for energy types supplied through networks (electricity, natural gas, district heating). This type of measured or metered data are rarely available for other types of carriers (e.g., fuels for cars, heating oil, biomass, coal).

Engineering estimates versus measured data; when to use it?

Engineering estimates will tend to be preferred over billing analysis when (TecMarket Works et al. 2004):

- No pre-measure billing data is available, e.g., new construction.
- Expected impacts are too small to likely be observed in a billing analysis (e.g., less than 10% of total consumption); remember that a billing analysis comprised everything behind the meter.
- The policy or programme deals with particular types of actions, e.g., with industrial process improvements; and the impact is again not measurable (or would be too costly to measure).
- The programme (or policy) includes requirements about energy audits or alike before and/or after the energy efficiency projects. Which makes that data from engineering calculations are directly available for the evaluation.

Engineering estimates can be applied to a few or many participants. As an example, engineering calculations are used in all the Energy Efficiency Obligation (EEO) schemes in Europe (Labanca and Bertoldi, 2016), and these schemes have very large number of participants.

It should be noted that (TecMarket Works et al. 2004) is about guidelines for evaluations of programmes done by utilities in California. For these programmes, utilities can have an access to the energy bills of their customers. Using data such as energy bills for such evaluation purposes is allowed under some conditions. The situation can be very different in European countries. Member States' regulations about the access or collection of energy bills can be restrictive, making it difficult for evaluators to access energy bills, as they often need to contact the participants to do so (with no guarantee of answer or approval) (see feedback from example n°1 in Ireland). This can therefore be a major reason to choose estimation-based approaches.

[10] Sometimes sub-metering is already in place for other reasons than evaluating energy savings, for example to control production processes. In those cases, the only cost is the cost of analysing the data.

This approach is best applied to programmes that involve equipment retrofits or replacing failed equipment with efficient models. Engineering methods have also been applied to estimating the effects of more complex measures, such as the adoption of efficient new construction practices or HVAC equipment through the use of building simulation models. In these cases, the models are calibrated to baseline conditions concerning building size, equipment, construction, and occupancy. Average savings are then estimated by changing the model parameters that are affected by the policy or programme. In some cases it is desirable and cost-effective to supplement engineering methods with short-term or spot-monitoring of site conditions. The results are then used to calibrate key parameters in the energy savings algorithms. Commonly measured parameters include operating hours for lighting and HVAC equipment, wattage for lighting and HVAC equipment, and temperature and pressure for various refrigeration and fluid applications (Vreuls 2005).

Billing analysis will tend to be preferred when (TecMarket Works et al. 2004):

- Both ex-ante and ex-post-billing-data are available (and if possible for several years to provide long enough time series to get results statistically significant).
- Impacts can be expected to be observed in a billing analysis (e.g., at least 10% of total consumption, depending upon method used, cleanliness of billing data, and accuracy of measured variables in analysis).
- The analysis is of a policy or programme with large numbers of actions or participants, for which the effect is measurable.

Billing analysis can be applied to a few or many participants. The latter could become less time consuming and therefore less costly and more attractive if (anonymized) smart metering is available. As an example, many of the individual companies within the Dutch long term agreements use metering for calculating the savings they achieve (see Veum, 2018).

Based on a review of case studies, Vreuls (2005) developed the following table for common applications of engineering estimates and savings calculations that involve measurements.

Table 2. Common applications of engineering- and measuring methods to estimate energy savings (Vreuls 2005).

Method	**Typical Policy Measures**	**Typical End-Uses**
Engineering	Economic Incentives: tax-related measures and rebates Information programmes: labelling Energy Audits	Residential Lighting and Appliances Commercial Lighting Industrial Motors
Engineering with building simulation modelling	Regulation: building codes Economic Incentives: tax-related measures rebates Information programmes: labelling	Commercial New Construction Residential New Construction Commercial HVAC
Engineering with monitoring	Economic Incentives: rebates Energy Audits Voluntary Agreements Regulation: building codes and equipment standards	Residential HVAC Commercial Lighting Industrial Motors C&I HVAC and Refrigeration Plug-load equipment
Bill Analysis	Economic Incentives: rebates Voluntary Agreements	Residential Shell and Heating Commercial Lighting & HVAC Residential New Construction Commercial New Construction
End-Use Metering	Economic Incentives: rebates Voluntary Agreements	Residential New Construction Commercial HVAC & Chillers C&I Lighting Industrial Motor Systems (Compressed air, pumps)

It should also be noted that in practice, several types of methods can be used, either subsequently (e.g., deemed estimates first for regular monitoring, then billing analysis for ex-post evaluation, as shown in example n°1 about Ireland) or simultaneously (e.g., deemed estimates for actions in the residential sector and detailed engineering calculations for actions in the industry sector, as done in many of the EEO schemes). The reason for this being the often relatively large programmes or policies that need to be evaluated and the related availability (or often non-availability) of input data needed to calculate energy savings, co-benefits and cost-effectiveness.

Likewise, deemed savings can be based on data that were possibly obtained from measurements. Estimated energy savings can thus be partly based on data that were measured previously, under the assumption that the results from these measurements can be used for similar actions or situations. The qualification of “estimated” or “deemed” data is therefore used to make it explicit that the data is not specific of the action or situation evaluated, but comes from another source.

It is also possible to use sophisticated calculations to estimate energy savings for typical situations. And then to use these estimates for similar situations, without collecting all the data. The main issue for this type of approach is to analyse under which conditions results can be extrapolated from one case to another.

Table 3 summarizes the pros and cons of the two evaluation approaches.

Table 3. Summary of the pros and cons of the two evaluation approaches

Estimation-based approach	
Pros	**Cons**
• No time lag to provide results • Energy savings assessed directly as the effect of the energy efficiency action(s) (no difficulty to assess small energy savings) • Results that reflect energy performance improvements in standardised conditions (results not affected by the differences in user behaviours, or by external factors) • Flexible approach from simple calculations or deemed estimates for types of actions that can be standardised, up to sophisticated modelling for complex actions • Can be easily combined with a monitoring system, and can enable to automatize energy savings calculations • Possibility to improve reliability of modelling by using calibration on actual energy consumption • Possibility to use methods for sensitivity analysis to assess uncertainties (for example due to assumptions on some of the variables)	• Results that do not necessarily reflect actual changes in energy consumption • Difficulties to verify the baseline situation • Do not allow ex-post assessment of net energy savings (further investigations needed, e.g. through surveys of participants, for more details see Voswinkel et al. 2018) • Risk of errors in data self-reported by participants or stakeholders • The more simplified the calculations, the higher the uncertainties. • The more sophisticated the calculations, the more data needed, the more costly. • Uncertainties due to factors not taken into account in the calculations/modelling (cf. simplification of the reality) • Uncertainties due to the assumptions about user behaviours (cf. prebound and rebound effects) • Uncertainties due to the assumptions about the quality of the action and its implementation (cf. performance gaps) • Uncertainties due to the possible use of reference or default values for some variables

Measurement-based approach	
Pros	**Cons**
• Based on actual energy consumption (energy bills or other metered data), so reflecting the actual changes in energy consumption (taking into account all factors: behaviours, performance gaps, external factors) • Enable comparison between participants and control (or non-participants) group (especially for a direct assessment of net energy savings, for more details see Voswinkel et al. 2018)	• Time lag due to the measurement period after the implementation of the actions (and possibly further delays to collect energy bills and to process data) • Difficulties to separate the different factors explaining changes in energy consumption: data needed to control them might not be available or too costly or complex to collect • Sub-metering or specific measurements might be needed when energy savings are

• Roll out of smart meters could increase the availability of metered data • Possibility to use methods for sensitivity analysis to assess uncertainties (for example due to sampling)	too small compared to random variations of metered energy consumption • Frequent difficulties to collect data from energy bills (cf. regulations about privacy and personal data protection; participants not willing to share their data; utilities not willing or not allowed to share data, or not enough utility staff available to answer requests) • Frequent data limitations in terms of time series (results might reflect particular, i.e. non-representative, conditions) • Frequent data limitations in terms of sample size, that can lead to statistical uncertainties when extrapolating results from samples • Technical issues of handling large set of data

Common issues raised by the comparison of estimated and measured (or metered) energy savings

Several stakeholders interviewed for the EPATEE case studies emphasised the importance to compare results obtained with different evaluation methods or to perform plausibility checks. The table summarizes the examples found in the case studies about comparing different methods.

Table 4. Examples of use or comparison of different methods (Broc et al., 2018).

Examples of use or comparison of different methods	Cases where these examples are mentioned
Plausibility check of the overall results (comparison with trends in energy consumption, and/or comparison with previous periods)	Environmental support scheme (AT), EE programmes of Vienna (AT), Primes Energie (BE)
Comparison of surveys and econometric analysis to assess additionality	EEO scheme (DK)
Comparison of different statistical methods	“Future investments” programme
Comparison of engineering calculations and billing data/analysis	Better Energy Homes (IE), Renovation programmes (LT), Subsidy scheme for housing corporations (NL), Supplier Obligation (UK), Warm Front (UK)
Comparison of monitoring of energy efficiency indicators (top-down approach) and monitoring based on engineering estimates at project level	Multi-year agreement (NL)
Comparison of standardised laboratory tests and field measurements	Purchase tax on new cars (NL)
Comparison of different methods to normalise energy consumption for weather conditions	WAP (US)

The objectives of comparing the results obtained with different methods are most often to test the robustness of the results (e.g. in view of decision-making) or to get a better understanding of the results (when the methods enable different types of analysis, with different advantages and limitations).

Plausibility checks do not aim at quantifying the differences between the results obtained with different methods, but to verify if their results are in the same range or order of magnitude. Likewise for the comparison of surveys and econometric analysis used to estimate additionality.

The other types of comparisons often aim at quantifying the differences between the results from different methods, and, whenever possible, at analysing the reasons for these differences.

Difficulties to compare estimated and measured energy savings can then be encountered when evaluating any type of policy measure in any type of sector. Various factors can indeed lead to differences between estimated and measured energy savings. It can be challenging to separate their effects and assess their relative contributions in the differences observed.

Generally, the differences between estimated and measured energy savings can come from:

- differences between assumptions about energy-using **behaviours** (in the estimations) and actual behaviours (reflected in measured data): this issue can be tackled by applying correction factors (cf. **prebound and rebound effects**);
- differences between assumed and actual **operating conditions**: this issue can be tackled by promoting good maintenance practices or **normalizing** energy consumption (e.g. using normalized weather conditions);
- differences between assumed and actual **quality of installation** of the energy efficiency product: this issue can be tackled by enforcing quality requirements or applying correction factors (cf. **performance gap**);
- differences between assumed and actual **energy performance** of the energy efficiency product or system: this issue can be tackled by enforcing quality requirements or applying correction factors (cf. **performance gap**);
- **simplifying assumptions** used in the model or calculation of the estimated energy savings (e.g. about interactions between products or systems): this can be tackled by improving the model or calculation formula, or by calibrating the model or applying correction factors;
- **uncertainties** in the measured energy savings (e.g., measurement uncertainties, sampling bias): this can be tackled by verifying the validity of the data and method used (e.g. statistical tests).

In the following, these issues are discussed for the case of policy measures promoting energy efficiency in buildings, due to a growing literature in this field. As mentioned above, similar issues can arise for energy savings from policy measures in other sectors (e.g. see the case of standardised vs. actual energy performance of vehicles).

For a general review of factors for differences between estimated and measured energy savings from energy efficiency improvements in buildings, see for example McElroy and Rosenow (2018).

Rebound and prebound effects

The **direct rebound effect** can be defined as follows:
"*Improved energy efficiency for a particular energy service will decrease the effective price of that service and should therefore lead to an increase in consumption of that service. This will tend to offset the reduction in energy consumption provided by the efficiency improvement*" (Sorrell and Dimitropoulos, 2008).

In the case of space heating, this means for example that, after improving the energy efficiency of a building, occupants might increase the indoor temperature (temperature setpoint), heat more rooms or use heating for longer periods. This is also presented as **comfort taking**.

For a detailed discussion and analysis of rebound effect (in general, not limited to buildings), see for example (Sorrell, 2007).

The **prebound effect** can be defined as cases where, before implementing an energy efficiency action, end-users tend to consume less energy than estimated by engineering models (for more details, see for example Sunikka-Blank and Galvin, 2012). Like the rebound effect, it is related to the assumptions made about energy using behaviours. But it can also encompass other assumptions made in engineering estimates that will tend to overestimate the energy consumption before the energy efficiency intervention.

In the case of space heating, this means for example that, before the energy efficiency intervention, occupants might restrict their use of space heating (e.g. lower indoor temperature, less rooms heated or for shorter periods), because a higher use of space heating would be too costly. Prebound effects can also be due to models estimating the energy consumption before the energy efficiency intervention based on the construction period of the buildings, assuming that no renovation work has been done since their construction. While in practice, some renovation works or other improvements can have been done.

Rebound and prebound effects are directly reflected in measured energy savings, as these are based on actual energy consumption. However, the measured data of energy consumption alone do not enable to separate these effects from other changes.

Assessing rebound related to space heating require to monitor indoor temperature in the different rooms of the building. See for example, the EPATEE case study about Warm Front in England (Broc, 2018). It can also be assessed more qualitatively based on surveys.

Prebound effects are not related to a change but to the initial situation. It can then be quantified by comparing estimated and measured energy consumption before the energy efficiency intervention. Such analyses have been done for example by comparing energy consumption estimated through Energy Performance Certificates (i.e. simplified engineering modelling) and metered energy consumption. See e.g. (Laurent et al. 2013), (Majcen et al. 2013) or (Summerfield et al. 2018). These studies have shown that dwellings labelled in the least energy efficient classes (G to E) tend to consume less than expected (due to prebound effects). Whereas the most efficient dwellings (B and A) tend to consume more than expected (due to rebound effects).

Like for rebound effects, understanding the different reasons for prebound effects and separating them require further data consumption (e.g. about indoor temperature, rates of renovation works previously done in buildings per construction period).

For a discussion about the differences between prebound and rebound effects and an example of analysis, see for example (Galvin and Sunikka-Blank, 2016).

Using results from previous studies on the same types of buildings and occupants can enable to define **correction factors** to take rebound and prebound effects into account in engineering estimates (see the pragmatic approach presented later on). Prebound effects can also be taken into account by improving the **calibration** of engineering models.

Performance gaps

Performance gaps can be defined as cases where the observed energy performance of the energy efficiency action installed is lower than the expected energy performance. For example, this can be due to differences in operating conditions, quality issues like defects when installing the actions, or defaults or overestimations in the energy performance of the product as stated by manufacturers or estimated with standardised laboratory tests.

Similarly to prebound and rebound effects, performance gaps are directly reflected in measured energy savings, as they are based on actual energy consumption. Further analysis is then needed to identify the reasons of the performance gaps and separate them from other changes in energy consumption. The analysis can for example include on-site inspections to check the quality of installation or the actual energy performance of the products or systems.

For an analysis about sources of quality defects related to energy efficiency in buildings, see for example Alencastro et al. (2018).

Including quality requirements in the policy measure can also help mitigate the risks of performance gaps.

Sources of uncertainties

The table below presents the various sources of uncertainties in energy savings results that could be identified from the review of the EPATEE case studies.

Table 5. Sources of uncertainties identified in the EPATEE case studies (Broc et al., 2018).

Sources of uncertainties identified in the EPATEE case studies	Case studies where these types of uncertainties are mentioned
Errors in the data reported by participants of stakeholders	Environmental Support scheme (AT), Primes Energie (BE), EEO scheme (DK), White certificates (IT), Voluntary agreements (FI), Voluntary energy audits (FI), Multi-year agreements (NL), Supplier Obligations (UK)
Assumptions used for engineering calculations (about behaviours, possible performance gaps)	EE programmes of Vienna (AT), Primes Energie (BE), Energy renovation of public buildings (CR), EEO scheme (DK), Better Energy Homes (IE), Renovation programmes (LT), Supplier Obligations (UK), Warm Front (UK)
Differences between actual characteristics and	Primes Energie (BE), Energy renovation of public buildings

Sources of uncertainties identified in the EPATEE case studies	Case studies where these types of uncertainties are mentioned
reference values used (about baseline conditions, and/or energy performance of the actions)	(CR)
Use of default values	Individual heat metering (CR), EEO scheme (DK), EE Fund (DE), Subsidy scheme for housing corporations (NL)
Statistical uncertainties due extrapolation from samples	Individual heat metering (CR)
Bias in the answers from surveys	EEO scheme (DK), EE Fund (DE)
Differences between actual performance in real conditions of use and standard performance assessed with normalised tests (e.g. manufacturers' data)	Voluntary agreements (FI), Voluntary energy audits (FI), EE Fund (DE), Purchase tax on new cars (NL)
Differences in the project as initially reported and the project finally implemented	"Future Investments" programme (FR)
Differences in calculation methods used by energy auditors (for complex actions)	EE Fund (DE)
Bias in sampling and/or matching methods	Better Energy Homes (IE), White certificates (IT), Nordsyn
Complexity of industrial processes (for very particular but large projects)	White certificates (IT)
Data limitations (sample size, time series)	EEO scheme (DK), Warm Front (UK)

Uncertainties can be challenging to quantify. Especially when the calculation method used is complex, many parameters or interactions involved, or when the uncertainties on the input data are unknown.

Sensitivity analysis can be applied to assess uncertainties, using methods with various levels of sophistication and providing results from basic orders of magnitude to detailed statistical analysis. In general, sensitivity analysis can be defined as a method to investigate how variations in input data or related assumptions can affect the output of a calculation model or formula (here energy consumption or energy savings). And reciprocally how variations in the output of the model can be attributed to variations in input data or related assumptions.

A simple approach of sensitivity analysis can be for example to define realistic ranges of values for the main parameters of the energy savings calculations. These ranges can then be used to calculate energy savings taking all the extreme values that tend to lower the savings (giving the *minimum minimorum*) and to increase the savings (giving the *maximum maximorum*). Both values provide a broad estimate of uncertainty range for the savings.

More sophisticated approaches can estimate or stipulate probability functions for each parameter included in the energy savings calculations, and then run large numbers of computer-based simulations to obtain a probability function of the output (energy savings) that then enables a statistical estimate of the uncertainty range for the savings.

For more details about methods for sensitivity analysis, see for example (Saltelli et al. 2008). More specifically about uncertainty analysis related to energy performance of buildings, see for example (Tian et al. 2018).

A pragmatic approach to take into account differences between estimated and measured energy savings: the use of correction factors to improve engineering estimates

The review of the EPATEE case studies (Broc et al. 2018) has shown that engineering estimates (or deemed savings) and billing analysis (or direct measurements) can be complementary. For example, deemed savings can be used to provide results on an on-going basis, while billing analysis can help to improve the reliability of energy savings results.

A pragmatic approach can thus be to include standard correction factors in the deemed savings to take into account the factors analysed above (rebound and prebound effects, performance gaps, etc.). And then to revise or improve these correction factors based on the results from billing analysis.

A good example of this approach is shown in the EPATEE case study about the Supplier Obligation in UK (Rosenow 2017). The deemed energy savings (also named deemed scores) include so-called "in-use factors" to take into account:

- differences in performance in-situ compared to laboratory testing and imperfect installation;
- natural variations in the thermal performance of structural and fabric elements that cannot be fully determined by the assessment, e.g. the possibility that the un-insulated walls have different U-values than the standard assumptions and that U-value varies across different parts of the wall;
- comfort taking (rebound effect) by the household, where some households may choose to heat their homes to a higher temperature;
- the household failing to operate the product/system effectively.

These factors are defined per action type, taking into account the results from available studies and feedback from monitoring & verification. Examples of in-use factors can be found in the Measures Table published by Ofgem for the Energy Company Obligation.

Such in-use factors were also defined for the Green Deal scheme (see DECC, 2012).

Revisions of the in-use factors can be based on results from analysis done with NEED (National Energy Efficiency Data-framework). NEED combines data from existing sources (administrative and commercial) to provide insights into how energy is used and what the impact of energy efficiency actions are on gas and electricity consumption, for different types of properties and households. For more details about NEED, see for example (Gregory and Prime, 2012) and the website of the UK ministry: https://www.gov.uk/government/statistics/national-energy-efficiency-data-framework-need-report-summary-of-analysis-2016
See also the related presentation made at the 3rd European EPATEE workshop.

When using correction factors, a good practice is to define them, whenever possible, based on available evidence about situations as close as the ones to be evaluated (e.g. similar types of buildings, similar types of actions). This will enable to define factors reducing the risks of over- or under-estimations of energy savings.

When no evidence base is directly available, correction factors can be defined from the literature, using conservative assumptions (to take into account differences between the situations analysed in the literature and the situations to be evaluated). Using default values based on conservative assumptions can also be a way to encourage stakeholders to provide data to revise the correction factors.

Examples of studies or evaluations including a comparison between estimated and measured energy savings

Two examples based on EPATEE case studies are presented in the next sections of this document.

Two EPATEE experience sharing webinars were also dedicated to this topic of comparing estimated and measured energy savings (or verifying energy savings). They provide four complementary examples:

Webinar #3

- Case 1: Green Investment Scheme in the Czech Republic, presented by Dr. Michaela Valentova, Czech Technical University in Prague. See also (Valentova et al. 2018).
- Case 2: Kirklees Warm Zone Scheme in UK, presented by Pr. Andy Gouldson, Deputy Director of CCCEP (Centre for Climate Change Economics and Policy), Leeds University. See also (Webber et al. 2015).

Webinar #4

- Case 1: Energy savings from renovations in the Dutch non-profit housing sector, with a statistical analysis of the differences between results based on the Energy Performance Certificates, and results based on metered energy consumption. Presented by Prof. Laure Itard, Chair Building Energy Epidemiology, Delft University of Technology (Netherlands). See also (Filippidou et al. 2017).
- Case 2: Analysis of the differences between energy consumption from building energy stock models, Energy Performance Certificates, and the impact of efficiency measures using metered data from the UK National Energy Efficiency Data-framework (NEED). Presented by Dr. Alex Summerfiel, Energy Institute, University College of London (UK). See also (Summerfield et al. 2018).

The presentation files and recordings of the webinars can be found at the links shown above.

Other examples from the literature:

- (Beagon et al. 2018): analysis of metered gas consumption of Irish social housing with a difference-in-differences method to assess the impacts of retrofitting by comparing retrofitted dwellings with a control group of dwellings. These results were then compared energy consumption levels estimated from Energy Performance Certificates, which enabled to assess prebound effects per type of buildings.
- (Gram-Hanssen et al. 2012) and (Raynaud et al. 2016): analysis of metered electricity consumption before and after installation of heat pumps, and comparison with expected energy savings to investigate possible rebound effects (particularly from using heat pumps for air conditioning in summer). The study by Gram-Hanssen et al. (2012) was done in Denmark, and included also a survey of households, and in-depth analysis on a sub-sample of 12 households (including technical inspection and qualitative interviews). The study by Raynaud et al. (2016) was done in the South-east of France, and included also a survey of households (about house and household characteristics, as well as about heating and air conditioning behaviours).

Concrete example n°1: [Ireland] Better Energy Homes

Background

The Better Energy Homes (BEH) scheme (formerly Home Energy Saving scheme) aims at improving the energy efficiency of dwellings built before 2006 by providing homeowners with fix grants per type of energy efficiency actions. A grant represents about 30% of the total investment costs (for the period 2008-2015) and can be used to upgrade dwellings with energy efficiency actions that may include insulation upgrades (wall and/or roof), installation of high-efficiency boilers and/or improved heating controls. The scheme started in 2008, and is still on-going. For more details about this scheme, see (Broc, 2017).

The monitoring and evaluation of the BEH scheme is structured in two ways:

- a regular monitoring of each application for a grant (amount of grants approved, number and type of actions carried out, energy savings, CO2 emissions avoided and jobs supported), together with random controls (including on-site inspections);
- complementary ex-post studies: a survey of over 10 000 participants (Motherway and Halpin, 2010), a Cost-Benefit Analysis (Scheer and Motherway, 2011) and an ex-post impact evaluation about actions implemented in 2009 (Scheer et al., 2013 ; SEAI, 2013).

The regular monitoring estimates energy savings achieved by each action using engineering calculations. These estimates provide the results for the reporting about the scheme.

BEH is the biggest grant scheme for energy efficiency in Ireland, making that it is subject to a particular attention. The Irish Public Spending Code requires that every scheme beyond a given threshold of annual public expenses shall be reviewed periodically. To answer to this requirement, SEAI (Sustainable Energy Authority of Ireland) did a first ex-post evaluation in the form of a Cost-Benefit Analysis (CBA) in 2011. The energy savings used in this CBA were the results from the regular monitoring (engineering estimates). This provided a first evidence base for the discussions with the Ministry of Finance.

The assumptions made for these engineering estimates (e.g., about comfort taking) were identified as one of the major sources of uncertainty in the CBA. It was therefore decided to complement the evaluation by an ex-post evaluation of the energy savings based on a billing analysis (measurement-based approach). This was done in 2012.

In practice, meter readings are a combination of actual physical reads by a qualified person, phone reads provided by customers and estimated reads calculated based on a historical use algorithm applied on a per dwelling basis.

This Irish case was chosen as an example for this topical case study, since it includes an in-depth ex-post evaluation, providing a feedback about a comparison between estimation-based (engineering calculations) and measurement-based (billing analysis) approaches. This is the subject of the analysis below.

Evaluation into practice

Table 6 shows the methods used in this Irish case, with one column presenting the data types and evaluation method used for the regular review and reporting about the scheme, and the other column presenting the same information for the ex-post billing analysis done in 2012 about actions implemented in 2009 (see Scheer et al., 2013).

Table 6. Methods used in the Irish example.

Type of data (see Table 1)	**annual reporting** (monitoring)	**ex-post billing analysis** (Scheer et al. 2013)
1. # of participants	Directly **monitored from the application process** (18203 homes and 31454 actions in 2009; about 200 000 homes and 337 000 actions over 2009-2016)	Sample built from participants having actions installed in 2009: initial sample of 500 participants with gas-heated dwellings and a Building Energy Rating done before and after the works. 250 participants gave agreement for access to their energy bills. The usable sample was finally of 210 participants after removing cases with changes of occupants and outliers. Sample of 153,928 households for the non-participant group.
2. # of actions taken		Actions installed by the participants of the sample were known from their application file.
3a. energy consumption before 3b. energy consumption after	**Simplified engineering estimated** calculation based on **deemed savings** per type of action (see details below the table).	**Billing data** (gas consumption / gas-heated dwellings only) Based on bi-monthly data that can be actual physical reads by the gas network operators, phone reads provided by customers and estimated reads based on a historical use algorithm. All dwellings included in the study had at least two actual reads per annum. *Period of observation*: - energy consumption before: at least 2 years preceding dwelling upgrades done by participants (i.e. 2007-2008) - energy consumption after: 1 year after the works (i.e. 2010) *Treatment group*: sample of 210 households who took part in the scheme in 2009. Billing data collected from the gas network operator (once the households gave their consent). *Comparison group*: 153 928 households who did not participate in the scheme. Billing data provided as an anonymised dataset by the gas network operator.
4. Normalization	The simplified engineering calculations are made under **normalised weather conditions.**	The statistical approach used (**Difference-in-Differences**, see details below the table) is meant to **control for external factors**: the **double comparison** (before/after and participants/ comparison group) is meant to ensure that the evaluated energy savings are related to the improvements carried out by the homeowners, and not to other factors that might affect all gas users (for example, fluctuations in usage relating to price, in GDP, extreme weather conditions). This assumption was tested using the consumption data of 2007 and 2008 for both groups that showed no significant difference between both groups in the changes in consumption in 2008 compared to 2007.

Complementary details about the results from **annual reporting**:

- The deemed savings per action type (used for monitoring) are based on results for 1500 dwellings observed during the pilot phase of the programme (in 2008), using **engineering calculations** similar to the method for energy performance certificates (Building Energy Rating in Ireland). The definitions of the deemed savings also took into account **two adjustment factors** through conservative assumptions:

  - **rebound effect**: based on the comparison between modelled and metered energy consumption during the pilot phase, conservative values for rebound effect (also called comfort taking) were defined, depending on the type of dwelling;
  - **free-rider (or deadweight) effects**: a default value of 18% was adopted, based on the results from the evaluation of the Energy Efficiency Commitment in UK (2002-2005)

- The **baseline** energy consumption used to calculate the deemed savings is equivalent to a **stock average** (average standard energy consumption before the implementation of works, defined per dwelling type).

Complementary details about the approach used for the **in-depth ex-post evaluation**:

- The ex-post evaluation used a **quasi-experimental approach** for the billing analysis, i.e. comparing a treatment group (sample of participants) and a comparison group (sample of non-participants). More specifically, the statistical approach used to evaluate the energy savings was the **difference-in-differences (DiD) method**, i.e. comparing the average change in the gas consumption over the observation period (i.e. between 2008 and 2010) in both groups (participants and comparison group). The difference between both average changes is thus assumed to be the energy savings resulting from the scheme. In other words, the **baseline** (or counterfactual) is the **change in energy consumption observed for the comparison group** between 2008 and 2010.
- For this evaluation, the treatment and comparison samples could not be selected randomly (which is a very common situation). So one challenge is then to **match** individuals (here dwellings) of both samples (participants and comparison group), to ensure that similar groups are compared and avoid (or limit) sampling bias. The variables used for this matching were the dwelling type and tenure (several possible matching variables were tested).
- Complementary data (in addition to gas bills) were thus needed to enable the matching and analyse the possible bias. For the sample of participants, surveys were used to collect the following data: dwelling characteristics, details of occupancy and some household behavioural data (in particular about comfort and use of secondary heating sources like a fireplace or plug-in electric heaters). For the comparison group, the dwelling type and number of bedrooms were directly available from the gas network operator's dataset for about half of the dwellings. Complementary sources were thus needed, such as national statistics or databases of Building Energy Rating (Irish Energy Performance Certificate) (for more details, see Scheer et al. 2013).
- If the test on 2007-2008 showed similar trends in both groups, it should be noted that the participants group had an average gas consumption (before intervention) significantly higher than the control group, mainly because the dwellings of the participants group were overall larger and older, therefore less energy efficient than the control group
- Data of the treatment and comparison groups were statistically cleaned to remove dwellings which were vacant and that had a change in the name on the gas account over the period of observation (2007–2010). In addition, outliers were removed (households whose energy consumption varied beyond specified minimum or maximum values).
- After that, average gas consumption indicators where calculated and the groups where compared with each other, using standard statistical tests.

- The in-depth ex-post evaluation also included a comparison between the energy savings estimated from the simplified engineering calculations and the energy savings measured from the billing analysis (see more details below).

Scheer et al. (2013) explained the choice of the difference-in-difference method compared to **other possible options**:

- **direct before–after comparisons** would have introduced bias due to significant changes in the environmental and economic conditions (significant reduction in economic activity and unusually cold weather in Ireland over the period of analysis);
- **cross-section estimators** require that the selection groups are statistically independent of the non-treatment outcome, while selection bias is likely among the participants to the scheme (for example due to higher environmental consciousness).

Lessons learnt (about the evaluation method/practice)

The ex-post evaluation presented by Scheer et al. (2013) is an interesting example that pointed several difficulties that can be encountered when using billing analysis to verify energy savings.

Difficulties in accessing / collecting billing data

A major issue was that for legal reasons, the evaluators (SEAI) had to contact participants to get their consent to use their billing data for the evaluation (even if these data were anonymised). This took a lot of efforts and led to a smaller sample compared to the initial plan.

The evaluators therefore strongly recommend to anticipate this issue when designing the scheme, and more specifically the process for households to apply for a grant. Requiring the consent at the time households apply for the grant (or validate the grant approval) can be an option to avoid later difficulties in collecting billing data from network operators, thereby saving significant data collection efforts.

The feasibility of this option might depend on the national legislations (e.g. about data privacy). This likely needs to be analysed by a legal team to ensure the conditions included in the scheme comply with the national legislation and cannot be contested afterwards.

Another issue encountered in collecting billing data from network operators is that their staff receives these requests for data on top of their regular workload. The evaluators thus mentioned that efforts were also needed to involve them in the evaluation.

Sampling and data limitations

Heating oil bills are not available on a regular basis, and does not necessarily correspond to the energy consumption between two deliveries (due to storage in tanks). Moreover, heating oil data are more difficult to collect (no network operator, various suppliers). Therefore, unmetered oil-heated homes were excluded from the study. Which creates a first important limitation for the representativeness of the sample.

The sample for the treatment group was further restricted to participants who had undertaken a before-and-after building energy rating (BER, Irish Energy Performance Certificate). This was made to ensure the availability of data on important variables, such as house type, size and specific energy characteristics (e.g. thermal properties). But this increased the risk of self-selection bias in the sample of participants: households who paid an additional fee for a before-and-after BER can be assumed to have a specific interest in the impact of the renovation works (e.g. on the value of the property). These households are also likely to have been more aware of their energy consumption before the works (compared to participants who were not interested in getting a BER).

Differences in the distribution of the actions per action type can also be observed between the participants group used for the ex-post evaluation and the whole participants.

There is thus a compromise between having more details/data about each individual in the sample, and having larger or more representative samples (see also balance between internal and external validity below).

Despite the efforts made to ensure data availability for the sample of participants, it was not possible to get data for all dwellings (in both samples) for some important explanatory variables. This created constraints in the method used, especially for matching. Other data sets providing information about parts of the samples (or basis for assumptions) were used to assess possible bias due to the limitations in the sampling and matching steps.

Difficulties to match different data sets / sources

Several possible sources of data were explored, to see how the billing data could be complemented with data on other variables to mitigate the risks of bias and improve the matching between both samples. However difficulties were encountered to match data sets at an individual dwelling level. Common identifiers could be used only for some parts of the samples. This is indeed a frequent problem when trying to combine large databases or results from different surveys. Especially when the data sources include confidentiality rules.

Risks of bias due to the conditions for the quasi-experimental approach

The treatment group could not be made by selecting randomly participants. Moreover, participation to the scheme is on a voluntary basis, inducing a risk of self-selection bias: the grants are open to all households, however certain profiles of households will be more willing to participate than others.

For example, the participants are more likely to be households that can afford the up-front capital needed for paying the works, and thus that can afford sufficient energy to comfortably heat their homes.

This was confirmed in the survey of participants done in 2010 (Motherway and Halpin, 2010), that showed for example a low participation rate of single adults and a high participation rate of retired households. Survey findings suggested that the socio-economic profile of participants is not comparable with that of the population.

Usual methods to correct for self-selection bias could not be used for this evaluation, due to the data limitations. It was indeed not possible to verify the similarity on both samples (treatment and comparison groups) for all key explanatory variables, inducing a risk of omitted variable bias.

However, the profiles of separate comparable data sets were used where possible to quantify the differences in the explanatory variables and how these might impact on the measured energy saving with reference to the relevant effects identified in the literature (in particular about rebound effects).

Balance between internal and external validity of the results

The matching of estimators can help mitigate the risk of bias, thereby improving the internal validity of the results (i.e. their statistical significance for the samples observed). However this reduces the external validity (i.e. the possibility to extrapolate the results to a wider population).

Similarly, the removal of outliers (i.e. dwellings with energy consumption above or below threshold extreme values) helps improving the internal validity, while reducing the external validity (unless further analysis can confirm that the removed outliers corresponded to data errors or very specific cases, and not to special categories that can also be found in the whole population).

Added value of the evaluation/study/project (for the policy and/or about the topic)

Comparison between the engineering estimates and the results from the billing analysis

The simplified engineering calculations estimated average final energy savings per dwelling of 33% (for the sample of participants considered in the ex-post study). The final ex-post evaluation showed average savings around 21% (±3%) compared to the comparison group (non-participants). This reduction of the savings by about 36% (±8%) could be due to the effects of behavioural changes (rebound effects), poorly performing equipment and potential inefficiencies in the systems installed, or limitations in the initial estimates of achievable savings (for example due assumptions included in the model or to the simplifying assumptions used to take into account technical interactions when several actions were implemented in the same dwelling).

The method used for the ex-post evaluation provides a statistical assessment of the uncertainty (cf. uncertainty range mentioned with the results above). However this does not take into account all the possible sources of uncertainty, such as the risks of sampling and matching bias explained above. Therefore part of the difference between estimated and metered energy savings can also be due to the limitations of the billing analysis.

The savings shortfall found in the ex-post evaluation was compared with the literature and stands in the same range as found in other similar studies.

The evaluators also pointed that these can be the different factors that can explain the difference between estimated and metered savings are hard to isolate or quantify.

Results that confirm a significant impact in terms of energy savings, but with limitations impeding extrapolations

Despite the savings shortfall mentioned above, the results of the ex-post evaluation confirmed that the energy savings impact of the actions supported by the scheme is substantial.

The results from this ex-post evaluation complemented the previous CBA (Cost-Benefit Analysis), by measuring the energy savings impacts and verifying assumptions that were identified in the CBA as key sources of uncertainty (for ex., assumptions about comfort taking). This made thus possible for SEAI to have more robust results and to be able to justify them to the Ministry of Finance. The results were very well received by the Ministry that was then interested in increasing the budget of the scheme.

But the evaluators also concluded that given the statistical differences between the participants group used in the ex-post evaluation and the population for dwelling and occupant characteristics that affect energy demand, it is not possible to draw conclusions from the ex-post evaluation that could be applied to the population of dwellings in Ireland (e.g. to estimate the potential for energy efficiency improvements). The sources of overestimation of the energy savings (in the participants group vs. the potential in the population) could be the lower initial energy efficiency and larger size of participants' dwellings. The sources of underestimation could be self-selection (higher energy awareness) and the main type of heating source (gas-heated dwellings having overall a lower consumption than oil-heated dwellings). The occupancy rate may also create a bias, potentially in both senses (participants include more working persons but also more retired households than the whole population).

Moreover, savings estimates were considered at a household (measure package) level, rather than by individual measures given the low frequency of uptake of some measures (in particular only 15 dwellings for external wall insulation and 20 dwellings for heating controls only). The results from the ex-post evaluation could therefore not be used directly to be the basis for standardised energy savings per action type. However they could be used to improve the standardised values (for example by updating the assumptions on rebound effect).

The evaluators concluded that a statistically significant sample, representative of dwellings in Ireland that have undergone an upgrade supported by the Better Energy Home scheme, could be more usefully compared to a comparison group of dwellings that had not undergone an upgrade. Such a study could provide sufficient data to assess the overall potential for improvement in Ireland's dwelling stock for a given suite of technologies. This is indeed the approach developed with NEED (National Energy Efficiency Data-framework) in UK.

For more details about this approach, see (Gregory and Prime, 2012) and the website of the UK ministry, or the related presentation made at the 3rd European EPATEE workshop.

Example of findings that helped to improve the scheme

In the interview done for the EPATEE case study, Jim Scheer highlighted that "*one of the key results of the evaluation was that the energy savings impacts differed depending on the initial level of energy performance of the dwellings (based on the Energy Performance Certificates' classes). We observed higher comfort taking in dwellings that were the least energy efficient before works. This argued in favour of promoting more comprehensive renovations*."

Direct feedback

Feedback from Jim Scheer (from the interview done for the EPATEE case study):

> **"***Empirical verifications represent a small budget compared to the whole budget of the scheme. Our experience with the ex-post impact evaluation is that it is really worth it.*

While engineering estimates are useful to monitor the results on an on-going basis, I strongly recommend to go beyond engineering estimates. One may have fear to do an ex-post impact evaluation, because it may show smaller results than based on the engineering estimates. However this increases the robustness of the results and therefore the confidence funders can have in them. This can be combined with a Cost-Benefit Analysis to show that despite the energy savings being possibly smaller, the overall result for society remains a net benefit, when taking into account all the impacts, beyond energy savings alone.

Qualitative analysis is also essential, for example to know how the participants feel about the improvements of their dwelling. This should be combined with the quantitative impact analysis, in order to understand how to promote the scheme."

For the complete feedback (e.g. about the reasons of doing the billing analysis, the outputs and the difficulties encountered), see the EPATEE case study: (Broc, 2017).

See also the presentation made by Jim Scheer at the first EPATEE experience sharing webinar: https://epatee.eu/events/webinar-1-part-1-how-energy-efficiency-policy-evaluation-can-produce-benefits-and-add-value

Concrete example n°2: [Lithuania] Renovation programme for multi-apartment blocks

Background

The Programme for the renovation (modernization) of **multi-apartment blocks** in Lithuania was adopted in 2004 and launched in 2005. The objective is to renovate the 4,000 most heat consuming multi-apartment blocks with 10.000 energy savings actions during the period from 2005 to 2020. The municipalities are in charge to evaluate and select multi-apartment buildings with the least energy efficiency that can then apply to the scheme. The process includes an **energy audit** and the definition of an investment plan. In most cases, a programme administrator appointed by the municipality is in charge of managing the financial support (combination of loans and rebates) and supervising the organization and quality of works, as well as monitoring of the results. The scheme also includes information actions to the building occupants about the actions installed and energy behaviours. For more details about this scheme, see (Skema and Dzenajaviciene, 2017).

The Lithuanian Energy Efficiency Law adopted in 2016 defines the policy measures to implement the Energy Efficiency Directive, as well as the roles and responsibilities about monitoring and energy savings calculations. It is thus **mandatory by law** that persons or organisations receiving financial support from public programs report the **entity's indicators** (heat consumption in kWh/m² per year) before and after implementing energy efficiency action(s).

This monitoring makes possible to **collect the following data**: type and number of actions implemented, energy properties of the actions, amount of investment for each action. It is relevant to program administrators who evaluate individual indicators, review the monitoring exercise, make the forecast and file the monitoring report on efficient consumption of energy for the previous calendar year to the Ministry of Energy.

For the period 2005-2016, almost 2,000 multi apartment blocks have been renovated (i.e. about 60.000 individual dwellings). Energy savings were estimated with detailed engineering calculations (see details below).

In parallel to the regular monitoring of the program, specific studies made **ex-post verifications** of energy savings on **samples of buildings**, based on measured data (Dagiliute and Luizyte 2011; Kveselis et al. 2017).

Evaluation into practice

Table 7 shows the methods used in this Lithuanian case, with one column presenting the data types and evaluation method used for the regular review and reporting about the scheme, and the other column presenting the same information for the ex-post studies (see Dagiliute and Luizyte 2011; Kveselis et al. 2017).

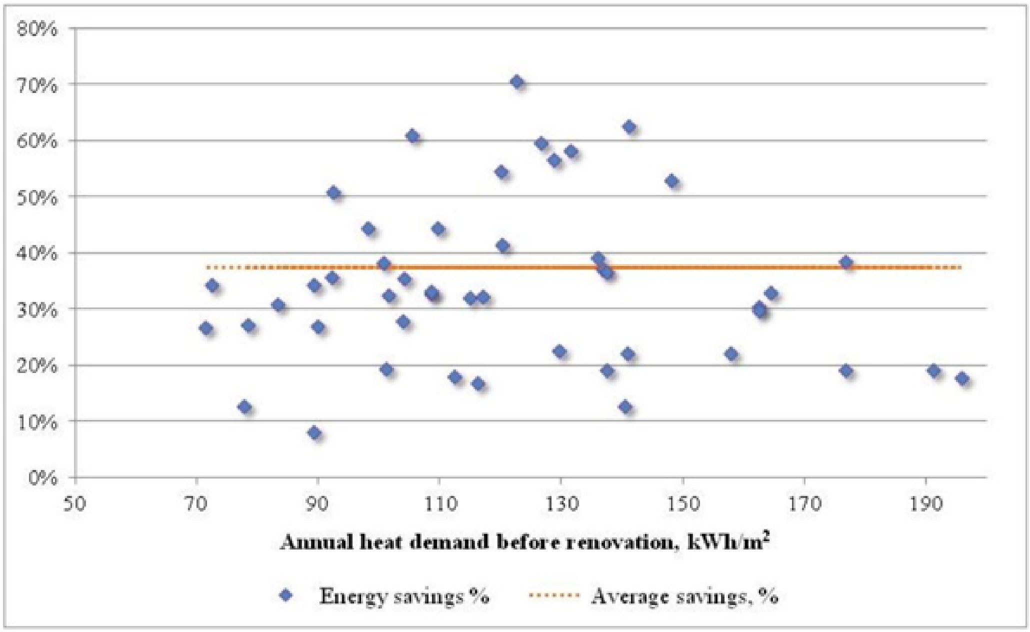
Table 7: Methods used in the Lithuanian example.

Type of data (see table 2)	**Annual reporting** (monitoring)	**Ex-post studies** (Dagiliute and Luizyte 2011; Kveselis et al. 2017)
1. # of participants	Mandatory monitoring of the project (1986 blocks buildings, i.e. about 59,580 dwellings, renovated over 2005-2016), that includes the details about the number and types of actions implemented.	Sample of 80 buildings equivalent to 2,400 dwellings (all in the same city).
2. # of actions taken		200 respondents were interviewed. 100 respondents were from multi-apartment blocks and 100 from detached houses (Dagiliute and Luizyte 2011).
3a. energy consumption before	Energy consumption before and after renovation is calculated using the methodology defined for the energy certification for buildings (**engineering calculations**) (for details about the software, see below the table).	Energy consumption (from district heating) before and after renovation is based on **metered data**: meter readings for 103 buildings which were fully or partially renovated. After a pre-screening, 80 buildings were chosen for deeper analysis (Kveselis et al. 2017).
3b. energy consumption after		
4. Normalization	The simplified engineering calculations are made under normalised weather conditions, and assuming normalised energy behaviours.	The measured heat consumption data for space heating was normalised using actual annual heating degree days at indoor temperature 18°C, estimated using the calculator of the Lithuanian Energy Agency (see www.ena.lt) and standard heating season degree days for Kaunas region and limiting outside temperature of 10 °C – 3789 degree days according to the construction norm RSN 156-94 (Kveselis et al. 2017).

In addition to the information given in the table:

- The **software** used to calculate the building performance is named NRG3. This is the national software tool created according to **national energy certification standard** STR 2.01.02:2016 'Design and certification of energy performance of buildings'. This software uses assumptions for normalised heating behaviours and weather conditions.

- **Energy savings based on engineering calculations** correspond to the changes in energy performance due to the actions implemented that are taken into account when modelling the energy consumption of the buildings after the renovation works (e.g. improved thermal coefficient of walls for insulation of walls). These energy savings have been estimated to be **on average at least 50%** (vs. estimated energy consumption before renovation works).
- The ex-post study based on **measured data** showed on average **32% energy savings** (average heat demand per square unit of building heated area before works: 145 kWh/m².year; after works: 97 kWh/m².year) (Kveselis et al. 2017). This would be 36% lower than the estimated savings (see also the spread of the results in Figure 1 below).
- The ex-post study analysed the spread of the energy savings rate among the buildings for which metered data could be collected. It found that there was no significant correlation between energy savings rates (in %) and building area nor with energy consumption before renovation works (see Figure 1 below), nor with investment per m². However, energy savings per m² of heated area were smaller in large buildings (> 4,000 m²) compared to small or mid-size buildings.



Note: Energy savings are here expressed in % of the energy consumption before renovation (including space heating and domestic hot water). The average energy savings is about 37%, with a spread from 9% to more than 70%. Considering space heating only, the average savings are about 32%.

Figure 1. Spread of the energy savings (in %) versus heat demand before renovation (in kWh/m².year) (source: Kveselis et al. 2017)

- There are **no data records on indoor conditions** before renovation. Higher indoor temperatures are observed after renovation, however exact figures are not known. Building owners stated that they have lower energy bills after renovation. They can consider higher comfort as affordable thus increasing indoor temperature from “normal” 20-21 °C to 22-24 °C. This can be one of the reasons for smaller measured energy savings compared to engineering estimates. For the same reasons, prebound effects would also need to be investigated to better understand the differences between estimated and measured energy savings.
- **Uncertainties** due to possible rebound effect or performance gap cannot be avoided even by conducting energy audits after renovation. Energy audits, however, can provide very useful information on renovation quality. Therefore such audits could be recommended on a case to case basis.

Lessons learnt (about the evaluation method/practice)

About **difficulties encountered** for the ex-post study, data to estimate annual energy savings per square meter of residential area was not as easily available as could be expected.

All multi-apartment residential buildings in Lithuania have installed heat meters for space heating and/or domestic hot water. Every apartment has also individual hot and cold water meters. Every resident can see in the bills monthly data for his/her apartment and the whole house. The space heating consumption for each apartment is derived from the metered value for the whole building divided by the apartment areas. However, this data is available for the residents of the apartments only.

The data the most commonly publicly available about energy consumption before and after renovation works are indeed estimates and not measured or metered data. Currently, the only mandatory action to investigate the effectiveness of building renovation is **energy certification of the buildings**. This is because renovations are almost all done thanks to public financial aids. And these aids require project holders to report the Energy Performance Certificates of the building or dwelling before and after the renovation works. However, this discloses the effectiveness of renovation works according to **standard assumptions** (no performance gap, normalised behaviours of the occupants, no prebound or rebound effect).

Some (measured) data is available publicly on the websites of a few district heating companies, and particularly of the Kaunas district heating company. This explains the geographical scope of the study. However, after renovation, part of this data was closed by the operators of the building heating systems. The reason is unknown. There is no legal regulation on these issues. Therefore, when there is no clear mandate for an evaluation, the ex-post studies are limited to the data that can be collected through direct contacts, particularly with the district heating companies. This was for example the case for this ex-post study (Kveselis et al. 2017). However not all companies are cooperative or interested in this type of study.

Measured data of energy consumption could be available from the databases of the district heating company, including data from monthly meters’ readings and data on building heating systems and substations. For the study (Kveselis et al. 2017), the research group was provided with measured heat consumption data for 103 renovated multi-apartment buildings in Kaunas City and Kaunas Regional municipalities. The data contained data on heated area, heat for domestic hot water, heat for space heating, total heat volume and annual heat consumption for space heating per square meter for the period of full 12 years (2004-2015). However, there was **no data on the year and type of renovation works** and implemented energy saving actions. In some cases, the renovation can be a simple change of substation or replacement of outside doors and windows. Pre-screening of the data was made to exclude from analysis the data where renovation year was not evident, i.e. when there was no visible effect seen from the data on heat demand for space heating. After this pre-screening, 80 buildings (more than 255,000 m^2 of heated area) were chosen for deeper analysis.

The data on heat consumption for space heating were distinguished in two sets of data, before and after renovation, for more evident renovation effect. However, as there are no exact dates from which we should account heat consumption as “after” renovation, there is some intermittent period considered as the period of the renovation process. There can be data included of the year of renovation, where data of first half of the year (January-April) contain metering readings of not renovated building and rest of the year contain data of already renovated building, as the renovation works usually take place during summer season when no space heating is required. In rare cases, building renovation could be performed in several steps and in different years. Such data rearrangement clearly shows cumulative renovation effect for all building stock and even provides some information on building performance after renovation.

Some increase in heat demand can be seen after 2-3 years which could be explained by customers' need for **better comfort** after they recognize that the cost for additional couple of degrees of indoor temperature is now **affordable** (which was not the case before the buildings were renovated). Nevertheless, this conclusion must be taken with caution, as most data cover only 1 to 3-year period after renovation. A longer time perspective would be needed to prove or reject such conclusion.
On the other hand, comfort level indicators, such as indoor temperature, relative moisture content and CO2 concentration, even fresh air flow were not investigated in the monitoring of the programme. Audits are not obligatory on adequacy of microclimate indoor conditions to sanitary regulations neither before nor after renovation. Certain conclusions can be gained from residents of the buildings or investment projects. However, they still cannot be considered as reliable data. Likewise, energy certification of buildings does not control the actual quality of the renovation works, and therefore do not provide information about possible performance gaps.

Based on the experience of this ex-post study (Kveselis et al. 2017), some **recommendations** can be drawn:

- For proper assessment of the renovation quality, one needs not only to perform energy certification of the buildings but also energy audit before and after renovation. There is a strong need for energy audit methodology to have uniform background for proper comparison of buildings before and after renovation.
- Audits should assess the quality of building envelope, heating and ventilating systems, as well as microclimate parameters (e.g. indoor temperature, humidity).
- Total actual measured heat demand in renovated buildings decreased by 32% only (vs. 50% estimated savings on average). This differs significantly from estimated values of energy consumption, which just confirms the need for more assessment. Also, the sample of buildings in the study done by Kveselis et al. (2017) might not be representative of the whole buildings renovated with the programme. This would thus be interesting to perform complementary ex-post studies in other cities and analysing if the results change depending on the type of buildings and occupants (e.g. income level).

Added value of the ex-post study (for the programme)

The study done by Kveselis et al. (2017) showed to the stakeholders that there is a **difference between measured and estimated data** (engineering estimates always showing at least 50% savings whereas the results measured on a sample of buildings show 32 % only on average). This shows that **more measured data is required** for proper assessment of renovation process. This also shows the need for such type of ex-post studies and to improve the availability of data needed for these studies. More demands from interested stakeholders (managers of the programme, ministries, responsible for implementation of the actions) would be needed.

At present, metering of heat consumption is possible only at the level of the whole building (not for each apartment). However, this already provides data about the actual renovation effectiveness. This requires also knowledge of **changes in parameters of the indoor environment**. This can be done during audit only and can be costly. But this difficulty could be overcome if this type of measurements (and audits) could be eligible for financial aids, as is currently the energy certification of buildings. This could be done on samples of dwellings, and would increase the knowledge about actual effectiveness of renovation works.

The results of the study thus shows that there is a need for changes in the policy, related to requirement of obligatory energy audits before and after renovation of residential buildings, as well as providing methodology for such energy audits, which would provide uniform background for methodologically based assessment. Depending on the size of the building (and thereby of the renovation project), special audits could be used, with a simplified but single methodology, with fixed number of necessary parameters. A **single methodology** is extremely important for comparability of results. Such energy audits could remain within reasonable cost levels, and would thus not be a barrier for renovation projects.

Besides there is also a need for more measurements of other parameters (especially about indoor environment) and more “open” data on estimated energy saving, which would enable proper assessment and comparison of energy savings.

The research team expects that main stakeholders, to whom the results as well as difficulties were presented, will make proper conclusions, based on the recommendations and will make respective changes in legislation, enabling better assessment of the results and will be more interested in measured data of costly renovation programme for large residential buildings, which make the majority of Lithuania’s residential stock.

A previous study (Dagiliute and Luizyte 2011) was focused on analysing the attitudes towards renovation (modernization) of buildings, and was done in the same area (Kaunas City). Interviews were done with occupants at their dwelling (100 dwellings/respondents). To find out consumers' **opinions and attitudes on building insulation**, the sample was built to have 50% of respondents living in multifamily buildings that were renovated, 50% of respondents living in detached houses. A statistical analysis was done to determine the **socio-economic factors** having an impact on consumer interest and choice of energy savings actions (renovation). The results of (Dagiliute and Luizyte 2011) suggested to rethink existing policy on multifamily building renovation; especially on its promotion and management, including revision of subsidizing schemes and minimizing bureaucracy. Whatever would be the energy efficiency instruments (here financial aids for building renovation), their efficiency depends much on the way the target groups are informed about benefits, obligations and so on.

Lack of information, hassle, disruption and high costs are the dominating barriers for energy efficiency improvements. Especially having in mind that building renovation is a highly costly (though cost-effective) action and disseminates rather slowly. Hence, wider spread of positive experiences (implemented projects, surveys) could be efficiently used for promotion and increased confidence in renovation programme.

Bibliography

General references

Alencastro, J., Fuertes, A., and P. de Wilde (2018). The relationship between quality defects and the thermal performance of buildings. *Renewable and Sustainable Energy Reviews*, 81, 883-894.

Beagon, P., Boland, F., and J. O'Donnell (2018). Quantitative evaluation of deep retrofitted social housing using metered gas data. Energy and Buildings, 170, 242-256.

Broc, J.S. (2018). Warm Front scheme (England). Case study prepared by IEECP for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

Broc, J.S., Oikonomou, V., and M. Dragovic (2018). Evaluation into Practice – Lessons learnt from 23 Evaluations of Energy Efficiency Policies. Report of task 3.2 of the EPATEE project.

Bukarica, V., Marić, L., and M. Matosović (2018). Individual heat metering in multi-family buildings in Croatia. Case study prepared by EIHP for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

DECC, 2012. How the Green Deal will reflect the in-situ performance of energy efficiency measures. Report of the Department of Energy and Climate Change, October 2012. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/48407/5505-how-the-green-deal-will-reflect-the-insitu-perfor.pdf

Di Santo, D. and E. Biele (2017). The Italian white certificates scheme. Case study prepared by FIRE for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

Dragovic, M., and J.S. Broc (2018). Nordsyn (Market Surveillance of Ecodesign and Energy Labelling) and the Effect Project in the Nordic Countries. Case study prepared by IEECP for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

Eichhammer, W. (2008). Distinction of energy efficiency improvement measures by type of appropriate evaluation method. Report of the EMEEES project, co-funded by the EU Intelligent Energy Europe programme. http://www.evaluate-energy-savings.eu/emeees/downloads/EMEEES_WP3_Report_Final.pdf

EPATEE (2017). Terminology and typologies used in the case studies. https://epatee.eu/case-studies

Filippidou, F., Nieboer, N., and H. Visscher, 2017. Effectiveness of energy renovations: a reassessment based on actual consumption savings. Proceedings of the ECEEE 2017 Summer Study, 1737-1746. http://proceedings.eceee.org/visabstrakt.php?event=7&doc=8-117-17

Galvin, R., and M. Sunikka-Blank (2016). Quantification of (p) rebound effects in retrofit policies–why does it matter? *Energy*, 95, 415-424.

Gram-Hanssen, K., Christensen, T. H., and P.E Petersen (2012). Air-to-air heat pumps in real-life use: Are potential savings achieved or are they transformed into increased comfort? *Energy and Buildings*, 53, 64-73.

Gregory, M., Prime, J. (2012). The development of a National Energy Efficiency Data-framework and its use in evaluating energy efficiency policies. Proceedings of IEPEC 2012. https://www.iepec.org/conf-docs/papers/2012PapersTOC/papers/033.pdf

Gynther, L. and U. Suomi (2017). Energy Efficiency Agreement for Industries in Finland. Case study prepared by Motiva for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

Labanca, N., and P. Bertoldi (2016). Energy Savings Calculation Methods under Article 7 of the Energy Efficiency Directive. Report of the Joint Research Centre for the European Commission. https://e3p.jrc.ec.europa.eu/publications/energy-savings-calculation-methods-under-article-7-energy-efficiency-directive

Laurent, M. H., Allibe, B., Oreszczyn, T., Hamilton, I., Galvin, R., and C. Tigchelaar (2013). Back to reality: How domestic energy efficiency policies in four European countries can be improved by using empirical data instead of normative calculation. Proceedings of the ECEEE 2013 Summer Study. https://www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2013/7-monitoring-and-evaluation/back-to-reality-how-domestic-energy-efficiency-policies-in-four-european-countries-can-be-improved-by-using-empirical-data-instead-of-normative-calculation/

Majcen, D., Itard, L. C. M., and H. Visscher (2013). Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications. *Energy policy*, 54, 125-136.

Marić, L., Matosović, M., and A. Kojakovic (2017). Energy Renovation Programme for Public Sector Buildings (2014 – 2015). Case study prepared by EIHP for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

McElroy, D.J., and J. Rosenow (2018). The performance gap of low-carbon building technologies and the role of public policy. Proceedings of IEPPEC 2018. http://www.ieppec.org/wp-content/uploads/2018/06/Wiese_paper_vienna.pdf

Meulen, J. and M. Menkveld (2018). Subsidy scheme for energy savings in social housing of Amsterdam. Case study prepared by ECN-TNO for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

Mort, D. (2017). Chapter 9: Metering CrossCutting Protocol. The Uniform Methods Project: Determining Energy Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), September 2017. https://www.nrel.gov/docs/fy17osti/68565.pdf

Raynaud, M., Osso, D., Bourges, B., Duplessis, B., and J. Adnot (2016). Evidence of an indirect rebound effect with reversible heat pumps: having air conditioning but not using it? Energy efficiency, 9(4), 847-860.

Rosenow, J., 2017. Supplier Obligations in Great Britain. Case study prepared by IEECP for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., ... and S. Tarantola (2008). Global sensitivity analysis: the primer. John Wiley & Sons. ISBN 978-0-470-05997-5.

SEEAction (2012). Energy efficiency programme impact evaluation guide. Prepared by Steven R. Schiller, for SEEAction (State and Local Energy Efficiency Action Network), December 2012. https://www4.eere.energy.gov/seeaction/system/files/documents/emv_ee_program_impact_guide_0.pdf

Sorrell, S. (2007). The Rebound Effect: an assessment of the evidence for economy-wide energy savings from improved energy efficiency. Report for UKERC, October 2007. http://www.ukerc.ac.uk/programmes/technology-and-policy-assessment/the-rebound-effect-report.html

Sorrell, S., Dimitropoulos, J., 2008. The rebound effect: Microeconomic definitions, limitations and extensions. *Ecological Economics*, 65(3), 636-649.

Spyridaki, N-A. and J.S. Broc (2018). Weatherization Assistance Programme (United States). Case study prepared by IEECP for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

SRC International A/S, et al., 2001. A European Ex-post Evaluation Guidebook for DSM and EE Service Programmes. Project co-funded by the European Commission's SAVE programme, April 2001. http://www.evaluate-energy-savings.eu/emeees/en/evaluation_tools/bottom-up.php

Summerfield, A., Oreszczyn, T., Palmer, J., and I.G. Hamilton (2018). Caveats for Policy Development when Combining Energy Ratings, National Building Energy Models, and Empirical Statistics. Proceedings of IEPPEC 2018. Available at: http://www.ieppec.org/wp-content/uploads/2018/05/Summerfield_paper_vienna.pdf

Sunikka-Blank M, Galvin R., 2012. Introducing the prebound effect: the gap between performance and actual energy consumption. *Building Research & Information*, 40(3), 260-273

TecMarket Works, Megdal & Associates, Architectural Energy Corporation, RLW Analytics, et al. (2004). The California Evaluation Framework. Report prepared the California Public Utilities Commission, September 2004. http://www.calmac.org/toolkitevaluator.asp

Tian, W., Heo, Y., de Wilde, P., Li, Z., Yan, D., Park, C. S., ... and G. Augenbroe (2018). A review of uncertainty analysis in building energy assessment. *Renewable and Sustainable Energy Reviews*, 93, 285-301.

Valentová M, Karásek J, Knápek J. Ex post evaluation of energy efficiency programs: Case study of Czech Green Investment Scheme. WIREs Energy and Environment. 2018;e323. https://doi.org/10.1002/wene.323

Veum, K.C., 2018. Long-Term Agreements on Energy Efficiency for the non-ETS sector (LTA3, the Netherlands). Case study prepared by ECN-TNO for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

Violette, D. M. and P. Rathbun (2017). Chapter 21: Estimating Net Savings - Common Practices. The Uniform Methods Projects: Methods for Determining Energy-Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), October 2017. https://www.nrel.gov/docs/fy17osti/68578.pdf

Vreuls, H. (2005). Evaluating energy efficiency policy measures & DSM programmes: Volume I – Evaluation guidebook. Prepared for the IEA-DSM Agreement Task 1-Subtask 9, October 2005. http://www.ieadsm.org/wp/files/Exco%20File%20Library/Key%20Publications/Volume1Total.pdf

Vreuls, H., Thomas, S. and J.S. Broc (2009). General bottom-up data collection, monitoring and calculation methods. Report of the EMEEES project, co-funded by the EU Intelligent Energy Europe programme. http://www.evaluate-energy-savings.eu/emeees/en/publications/reports/EMEEES_Bottom_up_draft_overview081006.pdf

Webber, P., Gouldson, A., and N. Kerr (2015). The impacts of household retrofit and domestic energy efficiency schemes: A large scale, ex post evaluation. Energy Policy, 84, 35-43.

References about example n°1 ([IRELAND] Better Energy Homes)

Broc, J.S. (2017). Better Energy Homes scheme (Ireland). Case study prepared by IEECP for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

Motherway, B. and T. Halpin (2010). Bringing Energy Home: Understanding how people think about energy in their homes. Report of the SEAI (Sustainable Energy Authority of Ireland), September 2010. http://www.seai.ie/News_Events/Press_Releases/Bringing_Energy_Home_Report.pdf

Scheer, J., Clancy, M., and S.N. Hógáin (2013). Quantification of energy savings from Ireland's Home Energy Saving scheme: an ex post billing analysis. *Energy Efficiency*, 6(1), 35-48. https://link.springer.com/article/10.1007/s12053-012-9164-8

Scheer, J. and B. Motherway (2011). Economic Analysis of Residential and Small-Business Energy Efficiency Improvements. Report of the SEAI, September 2011. http://www.seai.ie/Publications/Your_Home_Publications_/Energy_Efficiency/Economic_Analysis_of_Residential_and_Small-Business_Energy_Efficiency_Improvements.45851.shortcut.pdf

SEAI (2013). Better Energy Homes Scheme: Impact Report – Billing Analysis. Publication of SEAI.

References about example n°2 ([LITHUANIA] renovation programme for multi-apartment blocks 2005-2016)

Dagiliute, R. and S. Luizyte (2011). Blockhouse Renovation: Some Insights from Pilot Survey of Households in Kaunas City. Proceedings of the 8th International Conference "Environmental Engineering", May 19–20, 2011, Vilnius, Lithuania. http://leidykla.vgtu.lt/conferences/Enviro2011/Articles/3/735-739_Dagiliute_Luizyte.pdf

Kveselis, V., Dzenajaviciene, E.F., and A. Lisauskas (2017). Effectiveness of Residential Buildings Renovation on the Example of Kaunas City. Proceedings of the 10th International Conference "Environmental Engineering", Vilnius Gediminas Technical University, Lithuania. 27-28 April, 2017. http://enviro.vgtu.lt/index.php/enviro2017/2017/paper/viewFile/561/388

Skema, R. and F. Dzenajaviciene (2017). Lithuanian Renovation Programme with EU funding. Case study prepared by LEI for the EPATEE project, funded by the European Union's Horizon 2020 programme. https://epatee.eu/case-studies

How to cite this case study

Sipma, J., Broc, J.S., Skema, R. (2019). Comparing estimated versus measured energy savings. Topical case study of the EPATEE project, funded by the European Union's Horizon 2020 programme.