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

EPATEE

Evaluating grants or subsidies for energy efficiency in existing residential buildings using billing analysis

This specific guide can be applied to evaluate the savings due to grants or subsidies for saving actions in **existing residential buildings** using the method **billing analysis**. It includes guidance and explanations specific to this combination of types of policy measure, sector and method. As well as links to general guidance and explanations that can also apply to this combination.

CONTENT

1 | USE OF THE GUIDE–AUDIENCE, OBJECTIVES AND FOCUS

The primary **audience** for this guide is energy efficiency programme designers, implementers or supervisors, and evaluators looking for guidance on the evaluation process of energy savings in the scope of this specific guide.

Although the application of the guide will generally concern the (sub)national level, account will be taken of issues at EU level when relevant (e.g. the specific format of saving figures for the EED, and more particularly its article 7).

About this previous step in the evaluation process (*the choice of the method*), see the guidance provided here. However, after presenting the capabilities and limitations of the guide at hand, the user will be offered alternatives for the method within this guide (see section 6).

The **objective** of this guide is to provide:

- Information on the scope of the guide that enables the user to decide whether it is suited to his/her needs, and whether complementary or additional method(s) could be needed or useful (section 2);
- Guidance about specifying the evaluation objectives and requirements (section 3);
- Guidance about key methodological choices to calculate energy savings (section 4);
- Guidance about the inputs (data requirements) and outputs of the method (energy savings metrics) (section 5)**;**
- Possible alternative methods (with pros and cons) (section 6)
- Background about evaluation results other than energy savings (section 7);
- Relevant examples, case studies and/or good practices (section 8);
- Relevant references for further reading (section 9).

The specific guide is intended for assessing realised (ex-post) energy savings. However, account is taken of earlier (ex-ante) evaluations of expected savings, if available (see section 4).

The **focus** of the guide is on impact evaluation, i.e. determining the energy savings, but not on how this has been reached through a step by step process with intermediate results (process evaluation).

Readers looking for the basic and general principles of energy efficiency evaluation may find the following link useful.

2 | SCOPE OF THE GUIDE – POLICY, SECTOR and METHOD

2.1 About grants or subsidies

Financial incentives to stimulate investments for saving actions can take **various forms**: direct non-refundable aids (grants, subsidies, tax credits or rebates, loans with low interest rates or financial guarantees on loans.

More information and examples on the different subtypes residing under the main type **(legislative / normative measures)** can be found here and here.

Analysis of financial incentives for energy efficiency in buildings can for example be found in (Federici et al., 2018; Olubunmi et al., 2016; Studer and Rieder, 2019).

The focus of this guidance is restricted to the most common types of financial incentives offered to households or other owners of residential buildings: grants and subsidies. However, most of this guidance can also be applicable to schemes using other types of financial incentives, but in practice adaptations might be needed (e.g. how to get relevant data from the incentive scheme).

The subsidy can be defined as a fixed amount or as a percentage of the investment costs. The amount or rate can depend on the type of action or on the level of energy performance achieved (the higher the performance, the higher the incentive). The subsidy scheme may make a distinction between labour costs and costs for materials or products. The **scope of costs** taken into account is particularly important when assessing the cost-effectiveness of the incentive scheme (see section 7).

The subsidy can also depend on the income level or other criteria related to the dwelling owner or occupant. This is especially the case when the objectives of the incentive scheme include alleviating **energy poverty**. When the design of the scheme takes into account income levels or other social criteria such as employment, age etc., a separate evaluation for low-income dwellings might be needed (see cost-effectiveness in section 7) and additional evaluation objectives formulated (see section 3).

Subsidy schemes often represent **significant budget commitments** for the State, regional authorities or other public bodies that decide about funding the schemes. These schemes are often implemented by a mandated body (also named "*entrusted party*" or "*implementing public authority*" in the EED). This managing body usually monitors the scheme (e.g. number of participants, amounts of funding provided, number and types of actions funded) and reports to the government or other public institutions (e.g. Parliament, Court of Auditors) about the results of the scheme. **Cost-effectiveness indicators** can thus be given a high importance in the reporting needs.

More detailed information on the evaluation of grants or subsidies can be found here.

2.2 Evaluation for a combination of policy measure types

When subsidies are combined with other types of policy measures (e.g. tailored energy advice, training programmes for building professionals), it is assumed here that the overall savings due to energy efficiency improvements in buildings are mainly resulting from the subsidies. However in practice, the evaluation often encompasses the **combined savings effect** of all policy measures promoting the implementation of the types of actions in the scope of the evaluation.

This guidance is not capable of attributing part of the (overall) calculated savings to each of the policy measures (see also Double counting in the section on Gross to Net savings).

For discussions about **policy mix or comparisons** of different types of policy measures, see for example (Boza-Kiss et al. 2013).

2.3 Evaluation when combined with energy taxes

The calculated savings effect for subsidies will overlap with that of the energy tax. The energy tax indeed increases the cost-effectiveness of the energy saving actions from the point of view of the investors (most often the building owners). Therefore, **both policy measures**, the subsidies and the energy tax, **improve the economic conditions for the investments** in energy saving actions. The subsidies will be reducing the investment cost. While the energy tax will increase the economic benefit from the actions (by increasing the economic value of the energy savings), thereby reducing the payback time.

The guide is not capable of attributing part of the (overall) calculated savings to either the policy measures at hand or the energy tax. It would indeed be very difficult to find a comparable group of buildings and occupants who would not subject to a different energy tax, as energy taxes are usually applied nation-wide.

Distinguishing the effects of subsidies and energy taxes can be investigated with **modelling of investment behaviours**, by simulating scenarios with different policy packages. Calibrating this type of model usually requires data about the number of actions implemented in the country (or other area considered for the evaluation). See (Boonekamp, 2004).

This type of analysis can be complemented with surveys or modelling about the **willingness to pay** for energy saving actions. Such studies can be helpful for the design of the financial incentives (e.g. identifying the level of incentive that can trigger actions) and to provide a basis to evaluate the additionality of the financial incentives (or the free-rider effects) (see also section 4).

2.4 About existing residential buildings

This guide deals with energy saving actions for improving the energy efficiency of **existing dwellings**. It cannot cope with new dwellings as for billing analysis based on comparing “Before” and “After” energy consumption, no “Before” data are available for new dwellings. For other baselines the application of billing analysis on new dwellings is too rare to be taken into account here.

This guide covers both **individual houses and multifamily buildings**. In practice, the subsidy scheme can be focused on specific segments of the dwellings stock. For example to tailor the scheme in order to tackle specific barriers, such as decision processes in co-owned buildings (or condominiums) or split incentives in the case of the rental sector. In most cases, this does not affect the way gross energy savings can be evaluated. This can however affect the way to analyse additionality or free-rider effects, when evaluating additional or net energy savings (see also section 4).

In terms of end-uses, this guidance deals with the **thermal end-uses**, i.e. **HVAC** (heating, ventilation and air conditioning), and **domestic hot water**. Energy saving actions in the scope of this guidance thus includes actions improving the energy performance of the building envelope (e.g., wall insulation, roof insulation, windows replacement), heating and cooling systems (e.g., boiler replacement, hydraulic balancing of the heat distribution system), ventilation systems (e.g. heat recovery ventilation) or the system for domestic hot water.

It does not deal with actions on electrical end-uses, such as appliances and lighting (see Specific Guidance 14).

Subsidies for the energy renovation of residential buildings can concern **single actions** (e.g. wall/roof/floor insulation, boiler replacement) or **deep renovations** (i.e. aiming at improving the global energy performance of the buildings). In most cases, the subsidies are conditional upon minimum energy performance requirements. The incentive scheme may also require other conditions, such as quality label or installation by a qualified building professional.

In practice, billing analysis is preferred when evaluating policies or programmes promoting deep renovations (Agnew and Goldberg, 2017). Deemed savings and engineering methods are more frequently used when evaluating policies or programmes promoting or monitoring single actions. About these methods, see respectively the Specific Guidance 9 and 3. See also section 6 for a comparison of alternative methods.

Information on (sub)sectors defined in the Toolbox can be found here, chapter 2, p.17

2.5 Evaluation for cross-sector saving actions

This guide can also be applied to subsidy schemes promoting the renovation of tertiary buildings (service sector). One difference can be that the share of buildings with electric heating systems can be higher in the service sector, and that it can be more difficult to separate the consumption from HVAC and from other electric end-uses in tertiary buildings, in case consumption from all end-uses are billed from the same electricity meter. However, this is not the case when the buildings are operated with BAS (Building Automation System) that can provide detailed data of energy consumption together with data for other relevant variables such as temperature setpoint or patterns of use of the buildings (e.g. working days vs. week-ends). For more guidance about cases where data from BAS can be used, see for example (Romberger, J. 2017). And more generally for more guidance about cases when data can be collected through an energy management system, see for example (Stewart, 2017).

2.6 About billing analysis

Billing analysis, also sometimes called “consumption data analysis”, deals with the **statistical analysis** of **energy consumption data from billing records**, i.e. from the data registered by the energy meters used by utilities (either energy distributors or energy suppliers) to bill their customers.

This guide considers cases where billing records include data metered each month (monthly data) or every two months (bi-monthly data).

It should also be noted that billing analysis can only apply to **energy types that can be regularly metered at their point of use**. For example, actions in buildings using heating oil or biomass will usually not be possible to evaluate with billing analysis. Depending on the scope of the incentive scheme, this can be an important limitation.

For more details about data issues (e.g. different types of billing data, periodicity of metered data), see section 5.

Statistical analysis aims at separating the effects of the energy saving actions implemented from other factors of changes in energy consumption (e.g. energy prices). This means that billing analysis is recommended for cases where it is expected that the energy savings will be large enough compared to that of other, not controlled factors, so that the evaluation results can be statistically significant. This is one reason why billing analysis is commonly used to evaluate savings from deep renovations.

Obtaining statistically significant results also depend on the **sample size** or the **length of the time series** available for the data analysed (see section 5 for more details).

One key advantage of billing analysis is that it can directly take into account **so-called technical interactions (**between different energy savings actions, e.g. insulation of the building envelope and replacement of the heating system), **performance gaps** (e.g. due to defaults in the installation of the energy saving actions) **and rebound effects**. That is why it is assumed to reflect "actual" or "real" energy savings.

However, **it does not mean that billing analysis can distinguish these effects** in the results provided. Distinguishing the different effects requires additional data that is not always possible to collect (due to technical, economical or legal constraints). For more details about these factors (performance gaps, rebound effects), see section 4.

Depending on the method used for the billing analysis, it can also capture **free-rider and spill-over effects** to some extent. For more details, see section 4.

A key component of the approach used for the billing analysis is the choice of the control or comparison group (e.g. same buildings but before the energy saving actions; similar buildings with no actions supported by the incentive scheme). One major challenge of the billing analysis is to ensure that the comparison group and the participants group are comparable, and that their comparison enables to remove the effects other than the effects of the incentive scheme, in line with the evaluation objectives. For more details about the **choice of comparison group** and **matching methods**, see section 4 (and particularly, *Calculation baselines*).

Various methods can then be used to perform the billing analysis. They mostly differ according to the following features:

- Choice of comparison group(s): see section 4 (Calculation baselines).
- Time series of energy consumption data: from one year to several years for each period (before (pre-period) and after (post-period) participating in the incentive scheme).
- Specification of the periods compared: full calendar years (e.g. in line with the budget years of the scheme), or rolling periods (e.g. using rolling 12-month periods according to the month taken as participation date).
- Approach of data treatment (for more details, see for example the section 4 of Agnew and Goldberg, 2017):
  - **Two-stage approach**: energy consumption data are analysed at the level of each individual building, applying weather normalization, then comparing normalized consumption for the pre- and post-periods. This first step provides for each building the weather-normalised change in energy consumption. The second step is a cross-sectional analysis to compare changes in the two groups (participants and comparison groups), using regression models to distinguish what share of the weather-normalised change in energy consumption is correlated to the participation in the incentive scheme. Different regression models can be used according to the data available and the evaluation objectives.
  - **Pooled fixed-effects approach**: only the second step described above is performed, all the energy consumption data (for pre- and post-periods, and from participants and comparison groups) are included in a single regression model.

The choice of the method and corresponding features usually depends on the possibilities to form comparison groups (see section 4, *Calculation baselines*) and on the data availability (see section 5).

The ideal case would be to use long time series with a random assignment of building occupants (or owners) between the participants and the comparison group. However this is very rarely possible. Therefore other options have been developed to take into account the different practical constraints encountered. For more details about this, see for example (Agnew and Goldberg, 2017; Wade and Eyre, 2015).

General information about the various evaluation methods can be found be found here, table 1 and 2. This source also covers the combination of the method at hand with other methods, which will be dealt with below.

2.7 Complementary methods to determine total savings

Depending on the approach used, **billing analysis** can provide results in terms of **unitary energy savings** (where the unit is usually a building or a participant) or **total energy savings** (i.e. the energy savings for the whole participants). The latter requires billing data to be available for all the participants to the incentive scheme. This is rarely the case in Europe (see section 5 for more details).

Therefore, we consider here the case where the billing analysis can provide results in terms of unitary energy savings, or energy savings for a group of participants (participants sample). Then a **complementary method** is needed to obtain the number of actions/participants in order to calculate the **total energy savings**.

When dealing with subsidies, the number of participants will usually be directly monitored by the scheme, as part of the process to approve the applications for the incentive. However, depending on the scheme, data about the types of actions implemented by each participant is not necessarily registered on a systematic basis.

Another issue can be that due to regulation of data privacy, it might **not always be possible to match data from energy bills and data monitored by the scheme** (about participants). In that case, a survey of participants might be needed to collect information complementary to the energy bills (e.g. number and types of actions implemented per participant). For more guidance about **survey methods** and related issues, see for example (Baumgartner, 2017).

The data collection for the ex-post evaluation can then be done on a sample of participants. For more guidance about **sampling methods** and related issues, see for example (Khawaja et al., 2017).

The choice of the sampling method will depend on the data available (e.g. to form stratified samples of participants according to certain characteristics such as dwelling type, household size or income level). Together with the choice of comparison group, the choice of sampling method also has an influence on the **external validity** of the results, i.e. the extent to which the results obtained on the sample(s) can be extrapolated to the whole participants.

For practical examples, see the EPATEE case studies about the Better Energy Homes scheme in Ireland (Broc, 2017) or the Weatherization Assistance Program in the US (Spyridaki and Broc, 2018). More examples can also be found in section 8.

For a general presentation about complementary methods, see chapter 6 here.

2.8 Additional methods to increase reliability of the results

An additional method can be applied on top of billing analysis to improve the reliability of the evaluation results, complement the analysis of the data or improve the cost-effectiveness of the evaluation approach.

Billing analysis concerns overall energy consumption which can be influenced by many other factors than the effects of the incentive scheme. Data needed to investigate these other factors are not always readily available from the monitoring of the scheme or from other existing data sources. Another issue can arise when data about other factors cannot easily be matched with data of energy consumption or participation to the incentive scheme.

Therefore, complementary surveys can be needed to collect the missing data or data that can be matched with dataset already at hand. For more guidance about **survey methods** and related issues, see for example (Baumgartner, 2017).

Additional methods might also be needed when the evaluation objectives include possible free-rider effects (see section on adjustment factors).

Additional methods are also needed when the evaluation objectives are either to understand the reasons for differences between actual savings (as assessed with the billing analysis) and expected savings (usually assessed with engineering methods, or based on previous evaluations).

A first approach can then be to **compare results obtained with billing analysis and results obtained with engineering methods**. For more details about this, see the dedicated EPATEE topical case study (Sipma et al., 2019) and section 6 below.

A second approach can be to include the savings estimates from the engineering methods as a variable in the regression model used for the billing analysis. This type of model is called **statistically adjusted engineering regression** (or mix method). See for example section 4.3.2.3 in (Agnew and Goldberg, 2017).

A third approach can be to do **targeted additional data collection** designed to investigate specific issues.

For example, specific surveys or measurement campaigns can be done to assess **pre-bound and rebound effects** related to changes in heating behaviours (e.g. changes in setpoint temperature, in adapting heating at night, in the number of rooms heated). Depending on the policy objectives, this can also be analysed as improvements in thermal comfort (for more details, see for example Hong et al., 2009).

Another example can be on-site inspections on samples of participant buildings to identify causes of **performance gaps** (e.g. defaults in the installation of energy saving actions).

For a more general discussion about possible combinations of billing analysis with additional methods see chapter 6 here.

3 | EVALUATION OBJECTIVES and REQUIREMENTS

3.1 Meeting evaluation goals and ambition

Typical evaluation objectives of using billing analysis can be:

- Assessing **actual energy savings** (i.e. reflecting possible performance gaps and the energy behaviours of the occupants, including possible prebound or rebound effects (see section 4 for more details).
- Assessing energy savings from the point of view of the end-users.
- Assessing net energy savings (for more details, see section 4).
- Billing analysis is not the most appropriate method when the primary objective is to get savings estimate on an on-going basis (without time lag). It is less appropriate when the objective is to assess the energy performance improvement of the participants' buildings independently of the particular energy behaviours of their current occupants (see section 6 about alternative methods).

The following table shows whether this tool can be used to report on general evaluation goals or criteria. See also this document.

General types of evaluation goals or criteria	Level of ambition	Remarks
Calculation of realized energy savings from saving actions	High	Conditions of use of the chosen analysis method have to be checked with statistical tests (e.g. distribution of the residuals). Reliability of the results can also be checked with statistical tests. The larger the samples and the longer the time series, the better. Part of the uncertainties can be quantified (e.g. with confidence interval at given probability threshold). Part of the uncertainties will remained unknown (e.g. due to self-selection bias or omitted variables).
Calculation of energy savings attributed to the policy measure(s)	Medium to high	The feasibility to evaluate net energy savings firstly depends on the feasible options of comparison groups (see section 4). Complementary surveys can help consolidating the results when possible bias create uncertainties in the results from the method used for the billing analysis.
Cost-effectiveness of saving action (for end-users)	High	See remarks above about realized savings. Billing analysis is particularly appropriate to evaluate savings from the end-user perspective, as it takes into account occupants' energy behaviours and possible performance gaps. A frequent issue can be linked to the (non-)availability of the total investment costs, as the incentive scheme sometimes only monitor the amounts of subsidies granted. See section 7 for more details.
Cost-effectiveness of policy (government spending)	Medium to high	See remarks above about realized savings and savings attributed to the policy measure. Usually the amounts of public spending on subsidies are directly monitored by the scheme. The administration costs are often small compared to the amounts of subsidies but should still be considered. See section 7 for more details.
CO_2-emission reduction from saving actions	High	See remarks above about realized savings. As the savings are evaluated from billing data, the results are usually available per energy type. Which enables the conversion into CO_2 savings by using standard emission factors. The choice of the emission factor can raise debates when dealing with electric heating. See section 7 for more details.
CO_2-emission reduction attributed to the policy measure(s)	Medium to high	See remarks above about realized savings, savings attributed to the policy measure and CO_2-emission reduction from saving actions. See section 7 for more details.

3.2 Reporting expectations

Billing analysis can be used to provide results directly to the building occupants or owners, for example as part of a review for an energy performance contract or other type of contracts with energy savings guarantees. This case is more closely related to M&V than to policy evaluation. For more details about this, see for example the option C of the IPMVP (EVO, 2016), or (Jacobson, 2017).

This specific guide deals with the evaluation of energy savings from a subsidy scheme, i.e. a policy measure. In this case, billing analysis is used to provide results for a group of participants, and whenever possible for the whole group of participants.

Usually billing analysis when the evaluation objective is to assess **savings in terms of final energy**, as the data used are meter readings (i.e. data of final energy consumption).

However, it can also be used to assess primary energy savings or CO2 savings, provided that the corresponding conversion or emission factors are available (see respectively sections 5 and 7).

In most cases, the billing analysis will be performed over time series enabling to evaluate **first-year savings**. Using billing analysis to assess directly lifetime-cumulated savings is not possible in practice at this moment in time, as this would mean to monitor billing data over 20 years or more (due to the typical lifetimes of actions done for renovating dwellings). However, digitalization in the future may help this issue(For example, an online inventory of digitally recorded bills for a multiannual period).

In theory, it is possible to perform again billing analysis several years after the installation of the energy savings actions to assess savings persistence. However in practice, this is rarely done as this requires further resources and it can be difficult to get data about participants' energy consumption several years after their participation. Consequently, evaluating lifetime-cumulated savings or cumulative savings over several years requires additional assumptions (see *Energy savings over time* in section 5 for more details).

Billing analysis can be used to provide results in terms of **monetary energy savings**. This is indeed an important metric when the evaluation objectives are focused on cost-effectiveness indicators. If the indicators required to assess monetary energy savings over the action or investment lifetime, then assumptions about future trends in energy prices will be needed (see section 7).

Whatever, the metric(s) used to report the results, the **documentation** of the results is particularly important, so that they can be reviewed. It is also strongly recommended to **explain in the reporting the validity of the results and how they should be interpreted**, especially for readers who are not experts in statistics.

3.3 Time frame for evaluation

Billing analysis implies to distinguish three periods (for the data of energy consumption):

- The **pre-period**: period before the installation of the energy saving action(s).
- The **blackout period**: period of installation of the energy saving action(s), which can last several weeks in case of deep renovations.
- The **post-period**: period after the installation of the energy saving action(s).

The literature recommends to collect enough data points (i.e. meter readings) for the **pre- and post-periods to be at least 1 year long**, due to the seasonality of the energy consumption and also often due to the need to separate heating from other end-uses (especially in case of electric heating) (see for example Agnew and Goldberg, 2017).

Assuming a favourable case with meter readings available on a monthly basis, and renovation works done within a single month, the energy consumption data needed for the analysis cannot be available before 13 months after the renovation works.

Then difficulties can be encountered in **processing the data** to prepare clean datasets, for example: quality checks, matching datasets from different sources, removing cases with unsolved missing data, identifying outliers, etc. Finally time is also needed to **analyse the data**, for example: verifying if the conditions for using the chosen regression model(s) are met, comparing different matching methods, comparing different regression models, etc.

Altogether, this makes that the **time lag between the installations of the energy saving actions and getting the evaluation results** is usually **from 18 to 30 months**, as for example mentioned by Adam Bricknell in his presentation about the UK NEED (National Energy Efficiency Data-framework).

For more general guidance about evaluation planning, see also planning of evaluation in the link here.

3.4 Expertise needed for chosen method

The use of billing analysis requires expertise about the following:

- Expertise in **statistical methods**, especially on comparing data for groups of energy users;
- Expertise in methods for **weather normalization** (see *Normalization factors* in section 4).

Depending on the cases, complementary expertise can also be needed:

- Expertise in **sampling** methods.
- Expertise in **designing and performing surveys** (for data collection).
- Expertise in **legal issues about personal data and data privacy** (cf. GDPR – General Data Protection Regulation).

While these fields of expertise are needed on the side of the evaluation team, it is also recommended that the evaluation customers have basic knowledge about them. If this is not the case, it is then recommended to have a steering or scientific committee that can provide the evaluation customers with an independent review of the methods chosen and results obtained.

As pointed about *Reporting expectations* above, the interpretation of the results should be done carefully, taking into account the assumptions and possible limitations of the methods and data used. This requires a good documentation of the evaluation, as well as clear explanations to help readers with understanding the results and their background.

4 | KEY METHODOLOGICAL CHOICES FOR CALCULATION OF ENERGY SAVINGS

This section deals with key methodological choices to be considered when calculating energy savings: consistency between ex-ante and ex-post evaluation, baseline (including the choice of the comparison group), normalization and adjustment factors. These choices are important to document when reporting energy savings, to ensure the transparency of the results.

General principles of calculating realized savings using different methods can be found here and here

Billing analysis deals with the statistical analysis of energy consumption data from billing records, i.e. from the data registered by the energy meters used by utilities (either energy distributors or energy suppliers) to bill their customers (see About billing analysis in section 2).The specifications of the method(s) to use for the billing analysis are usually based on confronting:

- the most appropriate methods according to the evaluation objectives (and thereby the data needed to use these methods), and
- the practical constraints, in terms of data availability, time and resources available for the evaluation.

Depending on the cases, the specifications can be the result of an iterative process.

Data issues are further discussed in section 5.

4.1 Matching method with earlier ex-ante evaluation

From the viewpoint of methodological consistency and data availability using the same method in the ex-ante evaluation and in ex-post evaluation might be an obvious choice. However, for ex-ante evaluation only a few methods are usually considered, namely deemed savings, engineering estimate and stock modelling. By essence, **billing analysis cannot be used as such for ex-ante evaluation**.

However, previous ex-post evaluations that used billing analysis can be used as a basis to define deemed savings. If the same methodological choices are made in the previous and new ex-post evaluations, then combining the deemed savings based on previous billing analysis (for the ex-ante evaluation) and the new billing analysis (for the ex-post evaluation) can provide a consistent basis for comparisons.

In practice, a different method than the one(s) used for the ex-ante evaluation can be applied for the ex-post evaluation, depending on the evaluation objectives, timeframe and data available for the situation after implementing the actions. For possible combinations of methods applied ex-ante and ex-post, see chapter 7, calculation approaches here.

One typical case can be an ex-ante evaluation based on engineering estimates (possibly included in a building stock modelling), then complemented by an ex-post evaluation using billing analysis. For more details about the comparison of engineering methods and billing analysis, see the dedicated EPATEE topical case study (Sipma et al., 2019) and section 6 below.

If billing analysis does not provide an acceptable combination with the earlier applied ex-ante method it might be useful to select another method (see examples of alternative methods in section 6).

In practice, ex-ante and ex-post evaluations are applied consecutively. The ex-post evaluation builds on an ex-ante evaluation that makes use of data coming from previous ex-post evaluation or studies (e.g. about previous periods of the same incentive scheme, or about the same types of energy saving actions as the ones promoted by the new incentive scheme). These previous ex-post studies could have used another type of method as well.

4.2 Calculation baselines

Energy savings are defined in general as the difference between the actual situation and a reference situation without the saving actions (and without the policy measures that influence these saving actions). In case of saving actions the reference situation can be defined using various calculation baselines. For further background, see here.

When using a billing analysis, the **actual situation** is defined as the energy consumption (obtained from the billing data) of the **evaluated participants after the installation of the energy saving actions**. The evaluated participants are the participants who get energy saving actions installed with the support of the subsidy scheme and during the period under evaluation.

Several baseline options are then possible to define the **reference situation**:

- energy consumption of the same evaluated participants, but before the installation of the energy saving actions (option named “before/after comparison” or “actual before” in the EPATEE typology);
- energy consumption of a comparison group over the same period as the period monitored for the evaluated participants (option named “with/without” or “target/control group” in the EPATEE typology);
- energy consumption level set in minimum energy performance standards or building code (option named “Minimum efficiency standards”

In the **before/after comparison**, the participants group will act as its own comparison group in a time-dependent comparison. For the corresponding evaluation issues see next sections on **normalization factors, adjustment factors and gross/net savings.**

The **with/without comparison** is usually made over a duration covering both periods, before (pre-period) and after (post-period) the installation of the energy saving actions in the participants group. This comparison is thus time-dependent (before/after) and case-dependent (participants/comparison group). This type of comparison is named the **Difference-in-Differences** (**DiD**) method. The energy savings are calculated as the changes in energy consumption in the participants group minus the changes in energy consumption in the comparison group.

This comparison is assumed to **control for effects due to other factors than the participation in the incentive scheme.** For the corresponding evaluation issues see next sections 4.3, 4.4 and 4.5 on Normalisation factors, adjustment factors and gross and net energy savings. This comparison requires normalization factors, adjustment factors and gross/net savings.

When setting the baseline as the **minimum energy performance standards** (MEPS), **weather normalization** is needed so that the energy consumption from the billing data can be compared with the normalized values of the MEPS. This baseline can for example be chosen when the objective is to assess **energy savings additional to the current regulations**. Due to the limited application of this case billing analysis with this baseline option is not described further.

For a more general discussion about calculation baselines, see also here.

4.3 Normalization factors

The savings calculation from billing analysis provides a change in energy consumption that might need to be corrected for influences on energy consumption other than the saving actions. These so-called normalization factors can be weather conditions (with effect on consumption) and changes in energy using activities, such as occupation rate.

Depending on the baseline options chosen the following normalization factors can be applied in the evaluation.

The use of a before/after or with/without comparison is assumed to partly correct for the possible other influences than the saving actions. However, if the time series available are limited (e.g. less than 3 years for each, the pre- and post-period), the energy savings can then happen to be calculated for a particular period, not necessarily representative of typical weather conditions. Therefore in practice, the billing analysis will most often include **weather normalization**, either directly or to verify if the results are sensitive to it.

Weather normalization applies a correction to energy consumption data to take into account differences in outdoor temperature during the heating season, and possibly the cooling season (when the buildings have a cooling system)., the weather-dependent variables will usually be Heating Degree Days (HDD), and possibly Cooling Degree Days (CDD) (when the buildings have a cooling system). For more details about methods for weather normalization in billing analysis, see for example section 4.3.1 in (Agnew and Goldberg, 2017).

A better alternative is therefore to **use a with/without comparison**. In this case, the changes in energy consumption in the comparison group are assumed to control for external effects such as energy prices. However, as discussed in *Calculation baselines* above, unless a RCT method can be used, there will always remain an uncertainty about the extent to which the energy behaviours of the comparison group reflect the energy behaviours of the participants group in the absence of the subsidy scheme (for example due to the issue of self-selection bias).

Dealing with changes in **occupancy rates** is usually done by removing from the samples analysed (in both, participants and comparison groups) the dwellings with long vacancy period (i.e. several months) within the period taken into account in the evaluation. This can be done by detecting unusually low consumption. Ideally, dwellings with other types of significant changes in occupancy (e.g. change of the occupants, change in the household size, change in employment status) should be removed as well. But they are often more difficult to detect and can require additional data collection. When using a with/without comparison, and if the samples are large enough, a default option is to assume that these changes have evenly occurred in the participants and comparison groups.

As billing analysis calculate energy savings based on metered data, the calculated energy savings directly take into account possible prebound and rebound effects or performance gaps:

- **prebound effect**: cases where, before implementing energy saving actions, occupants tend to heat (or cool) less than could be assumed (i.e. compared to conventional energy behaviours, as for example defined in building codes). Prebound effect can for example occur in dwellings with very bad energy performance, making that the occupants cannot afford heating (or cooling).
- **rebound effect**: part of the improvement in the energy performance of the dwellings (and thereby of the decrease in the energy costs to heat or cool) might be used by the occupants to improve their thermal comfort (e.g. higher heating or lower cooling setpoint temperature), thereby decreasing the energy savings.

- **performance gap**: cases where the actual energy performance of the energy saving actions installed is lower than the expected energy performance, for example due to quality issues such as defects in the action installation, or due to inappropriate operating conditions.

The energy savings calculated with billing analysis do not need to be corrected for the effects listed above. However **it does not mean that billing analysis can distinguish these effects** in the results provided. Distinguishing the different effects requires additional data collection and analysis (see *Additional methods to increase reliability of the results*, at the end of section 2).

4.4 Adjustment factors

Adjustment factors define which part of the calculated energy savings can be attributed to a policy measure or meets the definition of savings specified in the evaluation objectives or reporting requirements (see below *Calculating Gross and net savings*). For a general discussion about adjustment factors, see here.

When using billing analysis, adjustment factors can concern the Free rider effect and the Spill-over/multiplier effect. **The need and relevance to use adjustment factors depend on the type of baseline (i.e. comparison group) chosen**.

Free rider effect deals with the energy saving actions that would have been installed by the building owners or occupants even in the absence of the incentive scheme. Such participants are commonly called free-riders or natural adopters. Energy savings from these actions should thus not be attributed, totally or partially, to the subsidy scheme. There can be several types or degrees of free-rider effect. For example, either exactly the same energy saving action would have been installed, or an energy saving action of the same type but with a lower energy performance.

Spill-over effect deals with the energy saving actions that would not have been installed in the absence of the incentive scheme, but that were not directly counted as part of the incentive scheme, because they were installed without receiving an incentive from the scheme. **Participants' spill-over effect** can for example be that some participants are installing other energy saving actions (e.g. replacing lightbulbs with more efficient ones) thanks to the bill savings achieved with the actions they have received incentives for (e.g. insulating the roof and walls). **Non-participants' spill-over effect** can for example be a homeowner who insulate the roof of her house after seeing the benefits her neighbor get from doing so, but who will not apply for the subsidy. In general, spill-over effect can be challenging to evaluate, especially because different types of spill-over effects can occur depending on the timeframe considered.

As discussed above about *Calculation baselines*, the ideal case would be to use a **RCT (Randomized Controlled Trials) method**. The random assignment of individuals to either the participants or the control groups is assumed to make that, statistically, both groups should include the same rate of natural adopters. Therefore, the comparison of the changes observed in both groups can be assumed to control for free-rider effects. Which means that **no further adjustment factor for free-rider effect is needed in this case**. This is also why a RCT method is meant to calculate directly net energy savings.

When it is not possible to use a RCT method (i.e. in most cases in practice), the comparison group used in the **quasi-experimental method** cannot be created on a purely random basis. Which means that it is very likely that the rate of natural adopters will be higher in the participants' group compared to the comparison group. Especially when the comparison group is formed with non-participants, it can be assumed that the rate of natural adopters among non-participants is very small, or even close to 0, as there is little reason why natural adopters would not apply for the incentive.

Therefore, **when using a quasi-experimental method**, **the share of free-rider effect** that would be controlled for by the comparison between the participants' and comparison groups **is unknown**. This means that the corresponding result is to some extent between gross and net energy savings.

When the comparison group is made of **earlier or later participants** (see above *Calculation baseline* for more details), it can be assumed that the comparison group includes no natural adopter. Therefore, the result from this approach can be considered **gross energy savings**. The advantage of using such a comparison group (e.g. vs. using a simple before/after comparison) is that it can control for changes due to external factors such as weather conditions and changes in energy prices.

In any case, when using a quasi-experimental method, if the evaluation objectives include the assessment of net energy savings, then an **additional method** (e.g. survey) is **needed to assess the free-rider effects**. For more guidance and examples about how to assess net energy savings, see the dedicated topical EPATEE case study (Voswinkel et al., 2018).

Whatever the type of comparison group chosen, the billing analysis can capture the participants' **spill-over effects** occurring during the post-period for which billing data have been collected. However, it cannot capture spill-over effect that would occur after this period.

Billing analysis cannot capture either nonparticipant spillover. And in case nonparticipant spillover effects would occur within the period observed for the evaluation (i.e. for which billing data have been collected), this can reduce the energy savings calculated with the billing analysis. As these effects will be included in the changes in the comparison group, and therefore included in the external effects deducted from the changes in the participants' group.

This means that **when an incentive scheme is successful to generate significant non-participants' spill-over effects on short term, then energy savings calculated with a billing analysis will likely underestimate the full net savings**. This risk of using billing analysis should be taken into account when choosing the evaluation method, depending on the evaluation scope and objectives, and after considering the likelihood of non-participants' spill-over effect on short term.

More generally, if the evaluation objectives include the **assessment of spill-over effects**, it is recommended to **use additional methods**. For example, considering methods to assess **market transformation effects**, and adopting a **long term perspective** (especially in case of long lasting incentive scheme).

When dealing with spill-over effects, the timeframe is a very important parameter. As pointed for example by (Agnew and Goldberg, 2017), non-participant spill-over from past policy measure (or previous period of the same incentive scheme) can affect the actions installed during the period under evaluation (by both, participants and non-participants). They indeed **influence the current market conditions**. Which will be included in the baseline. Therefore these current spill-over effects due to past policy measures will not be included in the energy savings calculated with billing analysis. Such issue should be taken into account when interpreting the results. Especially when the evaluation deals with an incentive scheme in place already for many years.

In parallel to free-rider and spill-over effects, in case of another policy focusing on the same saving actions as the evaluated incentive scheme, it might also be needed to consider corrections for **Double counting**. A possible way to track and correct for double counting is to put in place a centralized database that can be used to cross-check the actions counted for different policy measures. When a common type of unique identifier can be defined for the monitoring of each policy measures, then this cross-checking can identify the actions that would be included in the results of several policy measures.

Then allocation rules can be defined to either allocate the energy savings to one single policy measure according to a priority order, or distribute the energy savings among the different policy measure based on pre-defined allocation shares.

Investigating the interactions between several policy measures and quantifying their relative effects go beyond the scope of this guidance.

4.5 Calculating Gross and Net energy savings

Gross energy savings are energy savings calculated from the point of view of the final consumers, i.e. independently of whether the participants to the policy measure would have acted the same or differently in the absence of the policy measure.

Billing analysis provides results in terms of **gross energy savings** if the chosen baseline is based on a **before/after comparison**, or based on a with/without comparison with the **comparison group being formed with earlier or later participants** (see above *Calculation baselines*, for more details about the types of comparison groups).

The calculation of gross energy savings includes normalization factors (see above).

If the evaluation objective is to assess gross energy savings, then it is recommended whenever possible to use the with/without comparison with a comparison group being formed with earlier or later participants. Compared to the option of a simple before/after comparison, the comparison with earlier or later participants will better enable to control for external factors such as changes in energy prices.

Net energy savings are calculated from the point of view of the public authority that provides the budget for the incentive scheme. Therefore this calculation takes into account effects related to the causality or attribution of the actions or energy savings to the incentive scheme. Net energy savings are also the metric needed when assessing the cost-effectiveness of the scheme from a society point of view (see also section 7).

Billing analysis provides results in terms of net energy savings if the chosen baseline is based on a with/without comparison with the **comparison group being formed with a RCT (Randomized Controlled Trial) method**.

Note: when the evaluation objective is to assess net energy savings, then the baseline is also commonly called the counterfactual.

As discussed above about *Calculation baselines*, it is rarely possible in practice to use a RCT method when evaluating a subsidy scheme. The alternative is then to use a **quasi-experimental method**. However, as explained above about *Adjustment factors*, the fact that the comparison group is not defined on a purely random basis makes that the **share of free-rider effect captures by the billing analysis** is **unknown**. Therefore the savings calculated with a quasi-experimental method are **somewhere between gross and net savings**.

It is then therefore recommended to use an additional method (e.g. survey) to consolidate the assessment of net energy savings. For more guidance and examples about how to assess net energy savings, see the dedicated topical EPATEE case study (Voswinkel et al., 2018).

For practical examples about evaluating net energy savings from incentive schemes for energy efficiency in buildings, see section 8.

5 | INPUT AND OUTPUT

5.1 Main data requirements, data sources and collection technics

The table below describes the main data needs/requirements to perform a billing analysis, and corresponding data sources and collection technics. These data needs correspond to evaluations that would use as baseline option either a before/after comparison, or a with/without comparison (see *Calculation baselines* in section 4).

Calculation subject	Data requirements	Possible data sources and collection technics
Energy consumption	Monthly or bi-monthly data points should at least cover 12 months for each, the pre- and post-periods. Data points should be based as much as possible on meter readings (not on estimates)*.	Depending on the national legislation about billing data, they can either be collected from the utilities (energy distributors or suppliers) or might need to be collected directly from the end-users. In any case, consent from the end-users is usually needed to comply with GDPR (General Data Protection Regulation) or equivalent regulation on data privacy.
Weather normalization	Data about external temperature, and data about the base temperature Data about external temperature should be linked as much as possible to the location of the participants' buildings (e.g. thanks to ZIP code).	Data can usually be collected from the national weather services organisation. Utilities have sometimes their own weather data. It is important to use the same type of data source, and same rule of matching between weather and buildings dataset, for all buildings included in the evaluation.
Participants' data and data about the actions installed	The types of data needed depend on the analysis method chosen, and on the complementary analysis included in the evaluation objectives. It is important to identify the data needs from the scheme outset, to enable the preparation of the data collection.	Data are usually collected either directly from the monitoring system of the incentive scheme (e.g. through the application files of the participants), or from surveys done specifically for the evaluation.

*: *billing data are not necessarily based on meter readings. Depending on the utilities and national regulation on minimum requirements for billing data, billing data can include a mix of metered and estimated energy consumption. For example, data can be metered every six months, and the data points in-between can be estimates based on algorithms developed by the utility to predict energy consumption of its customers.*

More detailed guidance and practical advice about data requirements, collection and preparation can for example be found in section 4.5 (Agnew and Goldberg, 2017, pp.39-45).

5.2 Energy savings in final terms or in primary terms

Energy savings can be expressed in final terms or in primary terms. See definitions about primary and final energy here.

Billing analysis calculates **savings in final energy consumption**, as it starts from the billing data, i.e. data of final energy consumption.

The energy savings results can then be expressed in other metrics, when needed for other evaluation or reporting purposes, provided that the data needed to convert from one metric to the other is available. It is thus important to identify the needs to express the results in different metrics, so that data used in the energy savings calculations are documented enough to enable future conversions.

As billing analysis starts from billing data, the data of final energy savings are usually available per energy type (e.g. electricity, natural gas, district heat). **Primary energy savings** can then be calculated from the final energy savings by applying the primary factor for each energy type.

Primary factors can usually be found from national energy balances (for natural gas and electricity). Primary factors for district heat can sometimes also be available from national energy balances, or from annual reports of the district heat companies.

When primary energy savings are evaluated over several years (e.g. savings cumulated over the lifetime of the energy saving actions), assumptions are needed about the evolution of the primary factor for electricity over time, due to possible changes in the electricity mix (e.g. related to renewable targets).

A **particular attention** should be paid to possible differences between final and primary energy savings (other than related to the primary factor) when the energy saving actions include actions about **energy switching**, i.e. changes in the energy sources used for heating (or more rarely for cooling). Depending on the type of energy switching, the action can for example lead to savings in primary energy, but no savings in final energy, or even at the opposite increase in final energy consumption (e.g. switching from electric heating to biomass or gas boilers). Or vice versa.

5.3 Energy savings over time

Implemented saving actions in a year lead to savings over a number of consecutive years. E.g. a more efficient boiler can save gas over its lifetime of about 15 years, insulation over up to 60 years. For savings from behavioural changes (e.g. if the incentive scheme promotes energy display devices with tailored energy advice) might be not much longer than the period of the behavioural intervention (e.g. period over which tailored advice or feedback is provided to the occupants).

Energy savings can be calculated in different metrics in terms of time reference, for example: year-to-year, annual, cumulated annual, cumulative. See the definitions here.

The calculated **yearly savings** concern the savings of all new saving actions in that year. In this approach only data for the savings in the chosen year are needed. Yearly savings are also sometimes called **first-year savings** (for example in the US) or **new annual savings** (for example in the frame of the EED article 7). Billing analysis usually provides results in terms of yearly savings.

Adding up the yearly savings over a period, provided that the energy saving actions are still delivering savings, leads to **cumulative savings**. Billing analysis can provide results in terms of cumulative savings, however this would require to use longer time series than usual (collecting data of energy consumption for several years after the installation of the energy saving actions). Therefore in practice, further assumptions are needed about the lifetime of actions and the evolution of energy savings over time (cf. savings persistence). For more details about lifetime and savings persistence, see for example (Hoffman et al., 2015). Finally, savings from a saving action can be discounted and summed up over the lifetime of the action See link here.

Examples of **lifetime values** can be found in the following sources:

- CWA 15693:2007. Saving lifetimes of Energy Efficiency Improvement Measures in bottom-up calculations. CEN Workshop Agreement, April 2007.
- EN 15459:2017. Energy performance of buildings — Economic evaluation procedure for energy systems in buildings. CEN standard, June 2017. (see annex D).

Using values of lifetime per action type requires that the monitoring of the incentive scheme (or additional survey of participants) provides data about the number and types of action implemented by each participant. An alternative can be to define an average lifetime value that can be applied to the whole participants, assuming an average mix of action type or average action package.

When the energy savings are counted for the whole lifetime of the energy saving actions, then the result is often presented as **lifetime-cumulated savings**.

Lifetime-cumulated savings can also include the application of a **discount factor**. This discount factor can be used:

- For economic reasons: for example when a rewarding of the lifetime-cumulated energy savings is credited at once when the action is installed (while the energy savings are actually achieved over the lifetime of the action).
- For technical reasons: for example, to take into account that energy savings can decrease over time (e.g. for behavioural actions).

When using **technical discount factors**, they usually need to be differentiated per action type, as the changes in energy savings over time depend on the action type. Overall, there is limited evidence about decrease (or increase) of energy savings over time (for more details about this issue, see for example Hoffman et al., 2015).

A general default discount factor (i.e. uniform for all action types) can also be decided to reflect the risks related to the investment in the energy saving action, as perceived by the final customers or investors.

The **choice of the metric** depends on the policy or evaluation objectives and reporting needs. For example, when the main objective is to keep the evaluation simple, the choice can be yearly savings, as no further data or assumption is then needed. However, this will favour short-term actions against longer term action. Incentive schemes for energy efficiency in buildings usually aim at promoting actions with long lifetimes, such as wall/roof/floor insulation. Which can therefore support the choice of lifetime-cumulated savings.

The choice of the metric can also depend on reporting needs, for example in the context of EED article 7. In this framework, a particular type of cumulative savings is needed, counting the savings achieved within the obligation period (currently 2014-2020, and then 2021-2030).

6 | ALTERNATIVE FOR CHOSEN METHOD

6.1 Alternatives for the chosen method

Billing analysis is a commonly chosen method when the evaluation objective is more specifically to assess or verify the energy savings actually achieved or to assess the cost-effectiveness or efficiency of the scheme (see also in section 3 *Meeting evaluation goals and ambition*). Billing analyses are particularly used in case of subsidy schemes promoting deep renovations. In such cases, the types of actions implemented are not necessarily monitored. This makes that methods based on the kind of actions have been implemented (e.g. deemed savings or engineering methods) more difficult to use. Also, deep renovations imply that it can be important to consider the interactions between the different actions (e.g. insulation of building envelope and replacement of heating system).

One drawback of using billing analysis is the time lag between the installations of the actions and when the evaluation results can be available (see in section 3 *Time frame for evaluation*).

Deemed savings can then be appropriate when the objective is to provide stakeholders with visibility and to evaluate results from the monitoring system, without delay. Deemed savings can be used to evaluate savings from action types that can be described in a standardised way. They are then a cost-effective way to assess savings from large numbers of similar actions (for each standardised action type). They are less relevant for energy saving projects that would be very specific, thereby requiring case-by-case calculations (e.g. deep renovations of heterogeneous buildings). For more details, see Specific Guidance 9.

Engineering calculations can be used in case of large or specific energy saving projects, and can also provide results with limited time lag. Engineering calculations are for example often chosen to evaluate the savings from the subsidy scheme when they are used anyway as part of the process of the application for the subsidy (e.g. when an Energy Performance Certificate is required as part of the application file). In this case, there is no extra cost of using these data for the evaluation of the incentive scheme.

Direct measurements are more rarely used for the evaluation of subsidy schemes because they are rather costly to apply to all individual dwellings with saving actions. They can for example be used as part of specific verifications or on-site inspections to complement another method.

The table below presents the pros and cons of these methods commonly used for evaluating energy savings from incentive schemes for energy efficiency in buildings (see also in section 2 *Additional methods to increase reliability of the results*).

Type of method	Pros	Cons
Billing analysis	• Provide data about actual energy consumption / energy savings (capturing prebound and rebound effects, as well as performance gaps) • Can be used to evaluate ex-post net savings (depending on the type of comparison group chosen, or whether adjustment factors can be assessed with an additional method)	• Can only be used for ex-post evaluation • Frequent difficulties to collect billing data (unless data collection carefully planned and prepared in advance, e.g. collecting participants' consent when actions are installed) • Difficulties to get representative samples (cf. sampling bias + data losses along the evaluation process) • Delays in getting the result (at least one year to get the consumption after installing actions + time to process and analyse data) • Difficulties to find relevant control or comparison groups (when assessing net or additional savings)
Deemed savings	• Provide visibility to stakeholders • No delay in getting results from the monitoring system • Low running cost (once the set of deemed savings has been defined) • Calculations directly related to the energy saving actions installed	• Use limited to action types that can be described in a standardised way • Do not reflect the energy savings achieved for a given situation, but an average result for a population of actions • Can require significant preliminary efforts (if many different action types included in the scope of the incentive scheme) • Quality depending on the data available to define deemed savings • Possible gaps between deemed savings and actual savings (due to prebound and rebound effects, and to performance gaps) • Additional method needed to evaluate ex-post the additionality of the savings
Engineering calculations	• Can be used for almost all action types (including deep renovations) • Can enable to automatize energy savings calculations (through standardised formula for simple cases) • Can reflect the energy savings achieved for a given situation (specific calculations) • Limited delay in getting the results (calculations can be done before the actions are installed)	• Require to collect data for each case (so can be costly if data collected only for this purpose and for large numbers of actions / projects) • Possible gaps between engineering estimates and measured savings (see references below) • Additional method needed to evaluate ex-post the additionality of the savings

Direct measurements	• Provide data about actual energy consumption (for the baseline and/or for the situation with energy saving actions) or about actual values for key parameters (e.g. indoor temperature) • Can be used to assess performance gaps	• Can be costly if measurements only done for this purpose and for large numbers of actions • If sampling is used, attention should be paid to avoid sampling bias (if data are to be extrapolated) • Additional method needed to evaluate ex-post the additionality of the savings • Delay in installing the actions (if used to verify the baseline, then time needed to make the measurements, unless data are already available) • Delay in getting the results (if used to verify the situation with energy saving actions, then time needed to make measurements after the actions are installed + time to analyse the data)

About the comparison between engineering estimates and billing analysis, see also:

- The dedicated EPATEE topical case study: (Sipma et al., 2019)
- The experience sharing webinars about "How and what can we learn from verifying energy savings first estimated with engineering calculations?"

7 | ADDITIONAL EVALUATION RESULTS

7.1 Calculating avoided CO_2 emissions

Depending on the priority objectives of the incentive scheme, the evaluation objectives can include the assessment of the results in terms of CO_2 savings (i.e. avoided CO_2 emissions).

In practice, billing analysis first provides a result of final energy savings. Then avoided CO_2 emissions can be evaluated from the energy savings by applying emission factors. Four key aspects are to be considered when choosing the emission factor(s):

1. Emission factors vary according to the **energy type**, so the energy savings need to be recorded per energy type. This is usually the case for a billing analysis, as billing data from different energy types (e.g. electricity, natural gas, district heat) are usually available separately.
2. Emission factors for a given type of energy **can vary over time** (especially for **electricity**).
3. Emission factors can be:

    a) **Direct emission factors**: that take into account the emissions generated when producing the energy used;
    b) **Lifecycle emission factors**: that take into account all the emissions generated from the extraction of the energy resources up to the dismantling of the energy plant.

Due to the differences that the choice of emission factor(s) can induce, it is important to document what emission factor(s) has(have) been used.

Emission factors can for example be based on official national emission factors used for the national inventory of emissions of greenhouse gases.

The conversion of **electricity** savings into CO_2 savings is however a special case, depending on the national mix for electricity production. Several choices are indeed possible, for example:

- **Average emission factor**, calculated from the total annual emissions from electricity production (possibly taking into account national imports and exports) divided by the annual amount of electricity consumed: this is a simple approach, but that might not reflect the fact that end-uses can have different times of use and thus correspond to different load profiles (while the emission factor for electricity can vary significantly between base load and peak load).
- **Emission factors per type of end-use**: this requires more sophisticated calculations (e.g. by decomposing the national load curves per type of end-use) that will be meant to use emission factors reflecting the differences in time of use (e.g. daily, seasonally).

As the incentive schemes for energy efficiency in buildings mostly deal with heating (and possibly cooling), emission factors specific to electric heating (or cooling) may be more relevant.

The choice between the two options above will then depend on the national electricity mix (cf. emission factor varying significantly with time of production or not), and whether electric heating represents a high share of heating in the dwelling stock.

If the evaluation objectives aim at assessing lifetime-cumulated CO_2 savings, it can also be needed to define a scenario about the evolution of the national electricity mix over the period of calculation (e.g. taking into account the objectives of shares of electricity produced from renewable energy sources).

The avoided emission of **other greenhouse gasses** due to energy savings are not included here, as these emissions (and more specifically their reductions) are generally negligible compared to CO_2 for actions in the residential sector.

IPCC (Intergovernmental Panel on Climate Change) provides a **detailed database of peer-reviewed emission factors.**

7.2 Calculating cost-effectiveness

Cost-effectiveness is the ratio between costs to achieve energy savings and the amount of savings and possibly other benefits.

A distinction can be made according to the point of view adopted to assess cost-effectiveness:

- Cost-effectiveness for the end-user or participant (e.g. payback time)
- Cost-effectiveness from the point of view of the public authority funding the incentive (e.g. comparing different types of policy measures according to the public cost of energy savings)
- Cost-effectiveness for society at large (e.g. social net present value)

For more details about the different perspectives, see for example (Breitschopf et al., 2018).

The table below summarizes the main costs and benefits taken into account according to the perspective adopted.

Point of view	Costs	Benefits
Participants	• Part of the investments paid by the participants	• Gross energy savings (note: the financial aids received from the incentive scheme is taken into account in the fact that only the part of the investments paid by the participants is included in the calculation)
Public authorities	• Financial incentives paid to the participants • Administration costs of the scheme • Losses in tax revenues (due to additional energy savings)	• Increases in tax revenues (due to additional investments made in energy saving actions vs. baseline scenario) • Net energy savings (e.g. if the indicator calculated is about the public cost of energy savings)
Society	• Part of the investments paid by the participants (for additional actions only) • Financial incentives paid to the participants • Administration costs of the scheme	• Net energy savings

NOTE: the table above does not deal with non-energy impacts. Depending on the context and objectives of the incentive scheme, non-energy benefits can be larger than the benefits from energy savings. When assessing the cost-effectiveness of an incentive scheme from the society's point of view, it is therefore recommended to consider if it is relevant to include non-energy impacts in the scope of analysis (see *Calculating other co-benefits* below for more details).

For more details about assessing cost-effectiveness or cost-benefit analysis, see for example (Clinch and Healy 2001; Collins and Curtis, 2017). And more generally about assessing the effectiveness of incentive schemes, see for example (Stern et al., 1986; Studer and Rieder, 2019).

7.3 Calculating other Co-benefits

Possible co-benefits from saving energy can concern:

- Reduction of energy poverty
- Better indoor climate
- Extra employment
- Impact on economic activity
- Reduced dependency on (insecure) energy import

It should be noted that the impacts from incentive schemes on each of these aspects are usually positive, but can also be negative (e.g. on State budget or distributional effects). Therefore, it is in general more appropriate to speak about non-energy impacts.

For a general background about non-energy impacts, see here.

Analysis about how non-energy impacts can have an influence on the political support or rationale for an incentive scheme can be found for example in (Kerr et al., 2017; Rosenow, 2013).

8 | CONCRETE EXAMPLES

EPATEE case studies:

- 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.
- 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.
- Spyridaki, N-A. & Broc, J-S. (2018). Weatherization Assistance Programme (United States). Case study prepared by IEECP for the EPATEE project, funded by the European Union's Horizon 2020 programme.

Complementary references about these examples:

- Blasnik, M., Dalhoff, G., Carroll, D., Ucar, F. & Bausch, D. (2015). Evaluation of the Weatherization Assistance Program During Program Years 2009-2011 (American Recovery and Reinvestment Act Period): Energy Impacts for Single Family Homes. Report ORNL/TM-2014/582 of the Oak Ridge National Laboratory, March 2015.
- 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.

Other examples of evaluations using billing analysis:

- Adan, H. & Fuerst, F. (2016). Do energy efficiency measures really reduce household energy consumption? A difference-in-difference analysis. *Energy Efficiency*, 9(5), 1207–1219
- Beagon, P., Boland, F., & O'Donnell, J. (2018). Quantitative evaluation of deep retrofitted social housing using metered gas data. *Energy and Buildings*, 170, 242-256.
- Coyne, B., Lyons, S., & McCoy, D. (2018). The effects of home energy efficiency upgrades on social housing tenants: evidence from Ireland. *Energy Efficiency*, 11(8), 2077-2100.
- Filippini, M., & Zhang, L. (2019). Impacts of heat metering and efficiency retrofit policy on residential energy consumption in China. *Environmental Economics and Policy Studies*, 21(2), 203-216.
- Valentová M, Karásek J, & Knápek J. (2018). Ex post evaluation of energy efficiency programs: Case study of Czech Green Investment Scheme. *WIREs Energy and Environment*, 2018;e323.
- Webber, P., Gouldson, A., & Kerr, N. (2015). The impacts of household retrofit and domestic energy efficiency schemes: A large scale, ex post evaluation. *Energy Policy*, 84, 35-43. a

9 | FURTHER READING

General guidance on evaluations

- Agnew, K. & Goldberg, M. (2017). Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol. The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), November 2017. https://www.energy.gov/eere/about-us/ump-protocols
- Baumgartner, R. (2017). Chapter 12: Survey Design and Implementation for Estimating Gross Savings Cross-Cutting Protocol. The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), September 2017. https://www.energy.gov/eere/about-us/ump-protocols
- Breitschopf, B., Schlomann, B., and F. Voswinkel (2018). Identifying current knowledge, suggestions and conclusions from the literature. Report of task 3.1 of the EPATEE project.
- EVO (2016). Core Concepts – International Performance Measurement and Verification Protocol. Efficiency Valuation Organization, October 2016. https://evo-world.org
- Hoffman, I., Schiller, S., Todd, A., Billingsley, M., Goldman, C., Schwartz, L., 2015. Energy Savings Lifetimes and Persistence: Practices, Issues and Data. Technical Brief, Lawrence Berkeley National Laboratory, May 2015.
- Jacobson, D. (2017). Chapter 5: Residential Furnaces and Boilers Evaluation Protocol. The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), September 2017. https://www.energy.gov/eere/about-us/ump-protocols
- Khawaja, S., Rushton, J. & Keeling, J. (2017). Chapter 11: Sample Design Cross-Cutting Protocol. The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), September 2017. https://www.energy.gov/eere/about-us/ump-protocols
- Qiu, Y., & Patwardhan, A. (2018). Big Data and Residential Energy Efficiency Evaluation. *Current Sustainable/Renewable Energy Reports*, 5(1), 67-75.
- Romberger, J. (2017). Chapter 19: HVAC Controls (DDC/EMS/BAS) Evaluation Protocol. The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), September 2017. https://www.energy.gov/eere/about-us/ump-protocols
- Stewart, J. (2017). Chapter 24: Strategic Energy Management (SEM) Evaluation Protocol. The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), May 2017. https://www.energy.gov/eere/about-us/ump-protocols
- Stuart, E.A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science* 25(1):1–21.
- Violette, D.M. & Rathbun, P. (2017). Chapter 21: Estimating Net Savings – Common Practices. The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), October 2017. https://www.energy.gov/eere/about-us/ump-protocols

- Voswinkel, F., Broc, J.S., Breitschopf, B. & Schlomann, B. (2018). Evaluating net energy savings. Topical case study of the EPATEE project, funded by the European Union's Horizon 2020 programme.
- Wade, J. & Eyre, N. (2015). Energy efficiency evaluation: The evidence for real energy savings from energy efficiency programmes in the household sector. London: UK Energy Research Centre.

About financial incentives for energy efficiency in buildings

- Federici, A., Iorio, G. & Martini, C. (2018). Buildings and incentives schemes. ODYSSEE-MURE Policy brief, July 2018.
- Maki, A., Burns, R. J., Ha, L., & Rothman, A. J. (2016). Paying people to protect the environment: A meta-analysis of financial incentive interventions to promote environmental behaviors. *Journal of Environmental Psychology*, 47, 242-255.
- Olubunmi, O. A., Xia, P. B., & Skitmore, M. (2016). Green building incentives: A review. *Renewable and Sustainable Energy Reviews*, 59, 1611-1621.

About policy mix or comparisons of policy instruments for energy efficiency in buildings

- Boonekamp, P.G.M. (2007) Price elasticities, policy measures and actual development in household energy consumption—a bottom up analysis for the Netherlands. Energy Economics, 2007, Vol.29, issue2,133-157 (part of PhD thesis Improved methods to evaluate realised savings, 2005).
- Boza-Kiss, B., Moles-Grueso, S., & Urge-Vorsatz, D. (2013). Evaluating policy instruments to foster energy efficiency for the sustainable transformation of buildings. *Current Opinion in Environmental Sustainability*, 5(2), 163-176.
- Murphy, L. (2014). The policy instruments of European front-runners: effective for saving energy in existing dwellings? *Energy Efficiency*, 7(2), 285-301.

About issues related to the use of billing analysis

- Goldberg, M. L., Agnew, G.K., Fowlie, M., Train, K. & Smith, A.B. (2017). Not Just Another Pretty Formula: Practical Methods for Mitigating Self-Selection Bias in Billing Analysis Regressions. Proceedings of IEPEC 2017. www.iepec.org
- Randazzo, K.V., Ridge, R.S., Wayland, S., & Smith, B.A (2017). Comparison Groups for Whole Building Program Impact Evaluations: They are Harder and Easier Than You Think. Proceedings of IEPEC 2017. www.iepec.org

About the comparison between engineering estimates and billing analysis

- Filippidou, F., Nieboer, N., & Visscher, H. (2017). Effectiveness of energy renovations: a reassessment based on actual consumption savings. Proceedings of the ECEEE 2017 Summer Study, 1737-1746.
- Summerfield, A. J., Oreszczyn, T., Palmer, J., Hamilton, I. G., Li, F. G. N., Crawley, J., & Lowe, R. J. (2019). What do empirical findings reveal about modelled energy demand and energy ratings? Comparisons of gas consumption across the English residential sector. *Energy Policy*, 129, 997-1007.
- Summerfield, A.J., Oreszczyn T., Palmer, J., & Hamilton, I.G. (2018). Caveats for Policy Development when Combining Energy Ratings, National Building Energy Models, and Empirical Statistics. Proceedings of IEPPEC 2018.

Examples of studies about pre-bound or rebound effects, or improved comfort

- Hong, S. H., Gilbertson, J., Oreszczyn, T., Green, G., Ridley, I., & Warm Front Study Group (2009). A field study of thermal comfort in low-income dwellings in England before and after energy efficient refurbishment. *Building and Environment*, 44(6), 1228-1236.
- Milne, G., & Boardman, B. (2000). Making cold homes warmer: the effect of energy efficiency improvements in low-income homes. A report to the Energy Action Grants Agency Charitable Trust. *Energy policy*, 28(6-7), 411-424.
- Raynaud, M., Osso, D., Bourges, B., Duplessis, B., & Adnot, J. (2014). Evidence of an indirect rebound effect with air-to-air heat pump: to have and not to use? Proceeding of IEPPEC 2014.

About effectiveness, free-riders, etc.

- Clinch, J. P., & Healy, J. D. (2001). Cost-benefit analysis of domestic energy efficiency. *Energy Policy*, 29(2), 113-124.
- Collins, M., & Curtis, J. (2018). Willingness-to-pay and free-riding in a national energy efficiency retrofit grant scheme. *Energy policy*, 118, 211-220.
- Collins, M., & Curtis, J. (2017). Value for money in energy efficiency retrofits in Ireland: grant provider and grant recipients. *Applied Economics*, 49(51), 5245-5267.
- Olsthoorn, M., Schleich, J., Gassmann, X., & Faure, C. (2017). Free riding and rebates for residential energy efficiency upgrades: A multi-country contingent valuation experiment. *Energy Economics*, 68, 33-44.
- Shen, P., Betz, R., & Yong, S. K. (2018). Energy efficiency subsidies, additionality and incentive compatibility under hidden information. *Energy Efficiency*, 1-14 (online first).
- Stern, P. C., Aronson, E., Darley, J. M., Hill, D. H., Hirst, E., Kempton, W., & Wilbanks, T. J. (1986). The effectiveness of incentives for residential energy conservation. *Evaluation Review*, 10(2), 147-176.
- Studer, S., & Rieder, S. (2019). What Can Policy-Makers Do to Increase the Effectiveness of Building Renovation Subsidies? *Climate*, 7(2), 28.

Example about persistence of savings

- McCoy, D., & Kotsch, R. (2018). Why the energy efficiency gap is smaller than we think: quantifying heterogeneity and persistence in the returns to energy efficiency measures. Centre for Climate Change Economics and Policy Working Paper 340, November 2018.

About non-energy impacts:

- Coyne, B., Lyons, S., & McCoy, D. (2018). The effects of home energy efficiency upgrades on social housing tenants: evidence from Ireland. *Energy Efficiency*, 11(8), 2077-2100.
- Kerr, N., Gouldson, A., & Barrett, J. (2017). The rationale for energy efficiency policy: Assessing the recognition of the multiple benefits of energy efficiency retrofit policy. *Energy Policy*, 106, 212-221.
- Rosenow, J. (2013). The politics of the German CO 2-Building Rehabilitation Programme. *Energy Efficiency*, 6(2), 219-238.

Acknowledgments & Disclaimer

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 746265.