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

EPATEE

Evaluating energy savings from mandatory energy labelling for household appliances using billing analysis

This guide can be applied to evaluate the savings due to legislative/normative measures on energy efficiency household appliances using the billing analysis method. 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 program designers, implementers or supervisors, and evaluators looking for guidance on the evaluation process of energy savings in the scope of this 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).

This guide is not about the preceding step in the evaluation process, the choice of the method. About this previous step in the evaluation process, 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 this guide 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 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 mandatory energy labelling

Mandatory energy labelling is one of the sub-types included in the general category **"Legislative/Informative"** measures. More information and examples on the different subtypes residing under the main type '**legislative/normative'** measures can be found in: the Odyssee-Mure platform and in the EPATEE Knowledge Base.

This guidance is focused on the sub-type "mandatory energy labelling" due to the Energy Labelling Regulation (2017/1369/EU, earlier Energy Labelling Directive 2010/30/EU) that sets requirements in terms of energy labelling for energy-related products. The Energy Labelling regulation has already witnessed a few transformations, with a label reform underway for 2021 onwards; the labels will be rescaled to a simpler A through G, by eliminating the A+ signs (see for example European Commission, 2019).

For more details about the European policy framework for energy-related products, see the dedicated pages on the European Commission's website. And for more details about mandatory energy labelling, see for example (Ecofys et al. 2014; Hirayama et al., 2008; IEA 4E, 2016; Tholen et al., 2017; Waide et al., 2013; Wiel and McMahon, 2005).

More detailed information on the evaluation of legislative / normative measures can be found for example here.

2.2 Evaluation for a combination of policy measure types

Mandatory energy labelling and Minimum Energy Performance Standards (MEPS) are often implemented concurrently, as is the case in the EU. The effects of Energy Labelling regulations are thus closely tied to the Ecodesign directive and the related minimum energy performance requirements.

Usually the same agencies and bodies deal with assessing the effects of both the energy labelling and minimum energy performance requirements, assessing the overall effect. Thus, it can be very **hard to disentangle** the effects of each of the above policy measures. That's why in practice, ex-post evaluations assess the impacts of energy labelling and MEPS as a whole, without attributing savings to either labelling or MEPS (Corry Smith et al., 2016).

Providing guidance about methods to evaluate interactions from several policy measures are beyond the scope of this guidance, however, it is important to consider what overlaps can occur. A possible approach to separate the effects of energy labelling and MEPS is to use **econometric methods**, taking advantages of changes in the regulations over time. See for example (Bjerregaard and Møller, 2019) (see also *Adjustment factors* in section 4 of this guidance).

Overlap can also occur with **other types of policy measures**, such as:

- Subsidy scheme for early replacements of inefficient appliances (see Specific Guidance 14 about the case of subsidy schemes for household appliances) ;
- Subsidy scheme for replacements of appliances for low income households (as part of policy to tackle energy poverty) (see also Specific Guidance 14)
- Taxes on electricity prices (see below)
- Programmes for energy advice for households
- Other voluntary labelling schemes

For more details about the policies implemented by Member States for energy efficiency in appliances, see for example the MURE database or (Tholen et al., 2017).

2.3 Evaluation when combined with energy taxes

The calculated savings effect for mandatory energy labelling of household appliances will overlap with that of the energy taxes (particularly with electricity taxes). Increasing the energy prices makes appliances with higher energy efficiency more attractive as their lifecycle cost decreases (compared to less efficient appliances). Both instruments (energy labelling and energy taxes) thus act in the same direction: encouraging customers to buy more efficient appliances. Hence an overlap when evaluating their respective effects.

However, in practice, with the progressive reinforcements of the MEPS, the incremental difference in energy consumption between remaining energy classes becomes lower. This means that the incentive to save is much lower: due to the Ecodesign and other industry standards, the market is already forced to get rid of the most inefficient products. On the contrary, the ranges of prices of appliances from one end to the other of the energy scale is becoming tighter, as the standards are pushing out the least efficient devices. The price difference amongst high- and low-efficient devices has also diminished.

Bjerregaard and Møller (2019) have also pointed out that it happens that newer and more efficient appliances are discounted when first put on the market. Therefore, the first spur of buying the most efficient appliances might be also due to this incentive. However, such distortions are not large enough to justify the trend of buying more efficient appliances due to increased transparency through energy labelling.

One practical way to **avoid overlap in the energy savings counted for energy labelling and energy taxes** can thus be to consider that the savings from buying more efficient appliances are attributed to energy labelling. Whereas energy savings from an efficient use of the appliances are attributed to energy taxes (that can for example counteract possible rebound effects). In practice, **this can be done by calculating first the savings from energy labelling, and then deduct them from the total savings calculated for the energy taxes.**

2.4 About household appliances

Household appliances considered in this guidance are electrical appliances that are not fixed to the building, and not meant as the standard heating option or other HVAC (Heating, Ventilation and Air Conditioning) equipment. The appliances covered are washing machines, driers, fridges, TV and radio, ICT etc.

As of September 2019, mandatory energy labelling in European Union is applicable to 5 product groups of household appliances: 1) dishwashers, 2) washing machines and tumble driers, 3) refrigerating appliances, including wine storage fridges, 4) lamps, 5) electronic displays including televisions, 6) cooking appliances.

For more details, see the European product database for energy labelling.

For more details and updates about the list of energy-using products covered by the EcoDesign and Energy Labelling regulations, see the European Commission's webpage about energy efficient products.

Information on (sub)sectors defined in the Toolbox can be found in the EPATEE terminology, section 2, p.17.

2.5 Evaluation for cross-sector saving actions

This guide is also applicable to evaluate cross-sector savings, e.g. more efficient copy machines and/or PCs in offices, provided that the data needed are available for all relevant sectors.

2.6 About billing analysis

Billing analysis concerns total energy consumption of individual end-users. It can show a change in consumption after implementing one or more saving actions, which represents the **unitary savings** of the saving action(s) (the unit being usually a participant). Combined with the number of actions (see next section on complementary method) this provides total savings.

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.

Statistical analysis aims at separating the effects of the energy saving actions implemented from other factors of changes in energy consumption (e.g. weather conditions, 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 random variations**, so that the evaluation results can be statistically significant. It can thus be difficult to use billing analysis to evaluate energy savings from appliances in dwellings using electric heating or cooling (unless the heating or cooling consumption can be properly separated from the other electricity consumption). Likewise, it can be difficult to use billing analysis to evaluate energy savings from small appliances. See section 6 for a discussion about alternative methods.

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 effects such as performance gaps** (e.g. due to appliances having lower actual energy performance compared to manufacturers' statements) **and rebound effects** (e.g. using more lamps when efficient lightbulbs decrease the consumption of each lamp). 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.

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

The **quality of the billing analysis** and the **reliability of its results** are usually verified through statistical tests: first to check that the **conditions to use the chosen statistical** model are met (e.g. by looking at the distribution of the residuals), second to check the **statistical significance of the results**. These statistical tests provide first inputs to review the **internal validity of the results**, i.e. the extent to which the changes in the dependent variable (here the energy consumption) are caused by changes in the explanatory variables included in the statistical model (e.g. the implementation of the mandatory energy labelling, and possible changes in this regulation over time).

Moreover, in most cases, it is not possible to perform the billing analysis on data covering all households who bought an appliance in the period under evaluation (as mandatory labelling applies to the whole market). The data needed for the analysis are often available for a sample of households (see also below *Complementary methods to determine total savings*). In this case, the results obtained for the sample need to be extrapolated to the whole households who bought new appliances. Which raises issues related to the **external validity** of the results obtained from the billing analysis, i.e. the extent to which results obtained on a group of households can be extrapolated to other households.

As a general recommendation, whenever possible, **testing several methods can help increasing the robustness of the results** and **improving their interpretations**.

Information about the various evaluation methods can 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

Complementary methods are methods that are required, in addition to the primary selected method (here billing analysis), to calculate energy savings.

Billing analysis is meant to calculate unitary savings. These unitary savings should be multiplied by the number of actions to have the calculated total savings. The number of actions or participants can be obtained in various ways, see the link here, table 2 and 3 for more information.

For more information about methods to calculate unitary savings, see the the link here, table 2.

When evaluating energy savings related to household appliances, billing analysis is usually used to assess energy savings from actual changes in energy consumption. Depending on the data available and the sample size, this assessment can be done either globally for all households in the sample having bought an appliance, or the assessment can be done for sub-samples per energy class and type of appliances.

It is very unlikely that the billing analysis can be done for the whole population (either all households, or all households having bought a new appliance). Billing analysis can usually be performed on samples of households. Therefore, there is a need to extrapolate the results from these samples to the whole population taken into account when evaluating the policy.

A complementary method is then needed to assess the **total sales of appliances**, and the **market shares per energy class** (for the approach with sub-samples).

The complementary method can be to collect **market sales data**. Data about total sales of appliances (per type of appliances) can for example be collected from national statistics. However, data about sales per energy class are rarely easily available, and can be costly to get from companies specialised in market data. An alternative can be to use automated collection of data from the internet (**"crawling" technics**). See for example (Bennich et al., 2017; Enervee, 2014).

The complementary method can also be a method used to evaluate the **market transformation** effects of the mandatory energy labelling. For more discussions about the evaluation of market transformation effects, see Specific Guidance 7 about mandatory energy labelling for household appliances.

2.8 Additional methods to increase reliability of the results

An additional method can be applied in addition of billing analysis to improve the reliability of the evaluation results and/or the cost-effectiveness of the evaluation approach.

Billed energy consumption can be influenced by many other factors than the policy measure/saving actions at hand. Part of these can corrected for (see section 4, section on normalization). However, a careful check of the results, using an additional method, is needed to assure reliable results.

The additional method measurement can be applied for a selection of end-users, where the measurement concerns (as much as possible) the energy consumption that is relevant for the saving actions at stake (i.e. direct measurement of the energy consumption of the appliances targeted by the policy). For more details, see Specific Guidance 4.

Direct measurements are usually done for a limited number of end-users or appliances and not for every year. Moreover, direct measurements can be done on-site (measurement campaigns in samples of dwellings). Or they can be done in laboratories (e.g. laboratory testing as part of market surveillance). In this latter case, the measurements will provide data specific to the appliances, but for standardized conditions of use that may differ from "real-life" conditions. Billing analysis and direct measurement can thus complement each other in a cost-effective way.

About additional methods to investigate the effects of mandatory energy labelling on purchasing decisions, see the additional methods mentioned in Specific Guidance 7.

For possible combinations with an additional method see section 6 in this link.

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, 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).

At the opposite, 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), or that the expected savings are small compared to the whole electricity consumption as metered for the energy bills (see section 6 about alternative methods).

The table shows whether this guide 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	Limited to high	Depending on the extent of the savings compared to the whole metered energy consumption. As a rule of thumb, if the expected savings are lower than 10% of the metered electricity consumption, then using billing analysis might not be appropriate. Conditions of use of the chosen regression model 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.
Calculation of energy savings attributed to the policy measure(s)	Low	The feasibility to evaluate net energy savings with billing analysis firstly depends on the feasible options of comparison groups (see section 4). As mandatory energy labelling applies to the whole market, opportunities of relevant comparison groups can be rare. Therefore, alternative methods used to assess market transformation effects can be needed (see Specific Guidance 7 for more details).

Cost-effectiveness of saving action (for end-users)	Limited to high	Cost-effectiveness from the end-users' point of view does not necessarily require to know the number of actions. It can be assessed for typical cases (for example, ratios per energy class and type of appliance) from the sample(s) used for the billing analysis. 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. However, as mentioned above, for distinguishing the effects of the replacement of appliances from other factors influencing households' electricity consumption, these effects should be large enough compared to households' electricity consumption. See section 7 for more details about the specific issues related to cost-effectiveness assessment.
Cost-effectiveness of policy (government spending)	Low	Same comment as for energy savings attributed to the policy measure(s). Public costs can include costs of defining and updating the energy labels, of providing information for raising awareness about the labelling scheme, and costs of market surveillance (e.g., laboratory tests of samples of appliances, on-site inspections in samples of stores, monitoring of information displayed on retailers' websites). See section 7 for more details.
CO_2-emission reduction from saving actions	Low to Limited	Most appliances are consuming electricity, so the same type of energy. Which makes it simpler for assessing CO_2 savings from the energy savings. However, there can be difficulties related to the emission factor for electricity (see section 7). See also above the comments about calculation of energy savings.
CO_2-emission reduction attributed to the policy measure(s)	Low	Same comments as for energy savings attributed to the policy measure(s). Same comment as above about the possible difficulties related to the emission factor for electricity.

The evaluation objectives depend mostly on the needs and priorities of policy makers (e.g. public authorities in charge of designing, revising or enforcing the labelling schemes). Experience about evaluating mandatory energy labelling schemes shows that two of the most frequent key evaluation questions are: 1) how well the policies are implemented and received (process evaluation) ; and 2) what have the policies achieved (impact evaluation) (Corry Smith et al. 2016). This guidance is focused on impact evaluation, and more specifically the evaluation of energy savings. For general guidance about other types of evaluation, see the corresponding general guidance.

For more information on verification of actual energy savings and attribution/baseline/corrections, see section 4, and for cost-effectiveness and emission reduction see section 7.

3.2 Reporting expectations

Possible reporting options:

- Gross and Net savings (see section 4)
- Annual savings (no cumulative savings because no specific saving actions, with life times, are regarded, see section 5)
- Follow-up rate of suggested saving options to end-users, per type of saving action and given other policy measures focusing on the same end-users (see section 2).
- Market transformation due to more transparency of efficiency of appliances through e.g. labelling and information campaigns, and changes in stock of supply towards more efficient appliances.

If the evaluation objectives aim at assessing cumulative savings or market transformation effects, then complementary or additional methods will be needed. See section 5 about the issue of cumulative savings. See Specific Guidance 7 for more details about evaluating market transformation effects.

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 or on issues related to energy poverty.

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

The length of the period under evaluation is dependent on

- the active period of the policy measure;
- the need to monitor developments before the implementation of savings actions (in case of methods based on before/after saving actions); and
- the time needed to present (reliable enough) results or impacts that fit into the decision-making process. In some cases, the periodicity of evaluation can be set by law.

When dealing with household appliances, billing analysis usually implies to distinguish two periods (for the data of energy consumption):

- The **pre-period**: period before the installation of the new appliance(s).
- The **post-period**: period after the installation of the new appliance(s).

For household appliances, the time needed for installing the action (i.e. the new appliance) is very short (usually a matter of a few hours, or even less) and can therefore be neglected. However, in case the new appliance would be more complex to use compared to the replaced one, there might be a learning phase. This phase may not be representative of the future “regular” use of the new appliance.

The periods used in the billing analysis should be long enough to enable to take into account differences in usage over time when relevant (e.g. working days vs. week-ends; seasonality).

Moreover, the data points used for the billing analysis should be as much as possible based on metered data. Caution should be taken that billing data can also include estimates of energy consumption.

The time lag between the installation of the appliances and when the evaluation results can be available therefore depends on the minimum duration needed for the post-period. The time lag also includes the time needed to process and analyse the data. Which should not be neglected, especially in case of large samples.

When **planning** an evaluation using billing analysis, it is essential to take into account:

1. The **time lag** in getting the evaluation results: setting when the results are needed for the decision-making process will impact the **period(s) of the labelling scheme that can be included in the scope of evaluation**. This is particularly important to consider when features of the scheme have changed over time (e.g. updates in the energy classes or in the way to present the energy labels) and that the evaluation objectives are to look at the effects of these changes.
2. The **preparation of the data collection**: this should include ensuring the conditions to access billing data, but also identifying the other data needed (e.g. about the households) and defining how the different datasets collected can be matched.
3. The planning of evaluation activities concerns regular monitoring of energy consumption and factors that define consumption, intermediate check of (ex-ante) estimated (unitary) savings through measuring or surveys, intermediate evaluations to improve the policy implementation and the final evaluation and reporting. If the engineering estimate method is not combined with physical ex-post measurements for verification, the timeframe for evaluation is relatively low. It can already be performed after the finalizing the planning phase for the energy efficiency action. See also planning of evaluation in the link here.

3.4 Expertise needed for chosen method

The use of billing analysis requires expertise about **statistical methods**, and especially regression models.

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

- Expertise in **sampling** methods (see for example Khawaja et al., 2017).
- Expertise in **designing and performing surveys** (for data collection) (see for example Baumgartner, 2017).
- 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.

In section 3.2 on Reporting Expectations, 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.

About dealing with mandatory energy labelling, knowledge on consumer behaviour is needed, as is to set up and evaluate the various sub-measures that fall under the category "Legislative/Informative" measures.

Similarly, a good knowledge of the policy background for energy efficiency in appliances (e.g. about the Ecodesign and energy labelling regulations) is a prerequisite to analyse the impacts of mandatory energy labelling and possible interactions with other policy measures.

3.5 Boundaries for the evaluation

The scope of the evaluation will depend on the evaluation objective. The geographical boundary will depend on what sub-measure is being evaluated. For example, energy labeling schemes are EU-wide, whereas training and education as well as energy auditing or market surveillance are normally national measures.

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, 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

The method calculates the savings due to dedicated information on saving options to end-users. The method should not only take into account the actual implementation of saving actions but also the path from information to implementation, especially when the objective is to assess net energy savings.

In this guidance, billing analysis is used to determine unitary energy savings. As explained in section 2, a complementary method is needed to assess total energy savings.

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 this guide on ex-post evaluation might be an obvious choice. However, billing analysis is not applicable for ex-ante evaluation due to lack of observed data.

The billing analysis method for the ex-post evaluation can be combined with other methods for the ex-ante evaluation such as deemed savings, engineering estimate and stock modeling, depending on the evaluation objectives, timeframe and data available for the situation after implementing the actions.

Ex-post billing analysis can be matched with ex-ante deemed savings, where deemed savings are defined for each saving action (see Specific Guidance 14 about using deemed savings for household appliances). Engineering estimates (or scaled savings) are often used for more complex energy using systems like buildings, which are not regarded here. Ex-ante stock modeling can also give a benchmark, provided that data were available for the different types of appliances analysed in the ex-post evaluation. Moreover, stock modelling of appliances is usually combined with deemed savings based on the average differences in energy consumption between energy classes (for more details about stock modelling, see Specific Guidance 5).

In practice, ex-ante and ex-post evaluations are applied consecutively. An ex-ante evaluation using deemed savings can for example be further improved by checking the deemed savings through the ex-post billing analysis.

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: Before/after, With/without, Trend, Target/control group and Minimum efficiency standards ; see further in this internal EPATEE communication regarding an assessment of energy savings in the references of the EPATEE Knowledge Base here.

Total savings are found by multiplying the unitary savings per action by the number of actions. The chosen baseline can define the total savings through the unitary savings, but also through the number of actions.

For billing analysis, the applicable baseline option can be "**Before/after**", where "before" represents the situation before the new appliances are installed (for the same households). This option can for example be used when the objective is to assess **gross energy savings** (i.e. the energy savings as perceived by the end-users). This approach can then be combined with a complementary method enabling to assess the net effects of the policy (see below *Calculating Gross and Net energy savings*).

An alternative is the baseline option "**Target/control group**" that enables to correct for (almost) all other influences than the mandatory energy labelling, including new efficient appliances bought by households in the absence of the mandatory energy labelling. This option could be chosen when the objective is to assess **net energy savings** (i.e. the energy savings that can be attributed to the policy). But this baseline asks for a (sample wise) monitoring of a group of end-users that are both, comparable to the "participating" households AND not subject to the mandatory 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: Before/after, With/without, Trend, Target/control group and Minimum efficiency standards; see further here. Energy labelling of appliances is, in practice, **rarely possible** to find, as mandatory energy labelling applies to the whole country, and even to the whole European Union (due to the European regulation on energy labelling).

Moreover, when evaluating a regulation, it is rarely possible to randomly assign households to the participants' group (i.e. households subject to mandatory energy labelling) and to the control group (i.e. households not subject to mandatory energy labelling). So, evaluators will usually speak of a "comparison group", instead of a "control group".

Using a **comparison group** to define the baseline can still be relevant depending on the evaluation objectives. For example, if the objective is to take into account possible effects of changes in the electricity prices, a comparison group can be used as part of a Difference-in-Differences (DiD) analysis. This method can for example compare the changes in electricity consumption in a participants' group (i.e. households who have installed a new appliance) and a comparison group (i.e. households who have not changed their appliances) to **investigate possible factors of changes in electricity consumption** other than the replacement of an appliance by a more efficient one. However, as the comparison group will have to be made from households living in areas subject to the same mandatory energy labelling scheme, there will be a high risk of **self-selection bias**. This means that households in the comparison group and in the participants' group will very likely have systematic differences about their energy-related behaviours. For more details, see for example (Khawaja et al., 2017).

An alternative option can be to define the baseline from the **minimum energy performance requirements** set in the EcoDesign regulations for the types of appliances under evaluation. This option can for example be used when the objective is to separate the effects of the mandatory energy labelling from the effects of the minimum energy performance requirements.

The baseline options above deal with the step of calculating unitary energy savings. Then when the objective is to assess net energy savings, the complementary method used to assess the number of actions will also need to define a baseline or counterfactual. See below *Calculating Gross and Net energy savings*.

For a general discussion about the choice of baseline, see for example the evaluation guidebook of the IEA (including also case studies).

4.3 Normalization factors

With the chosen baseline a change in energy consumption can be calculated, but this change should be corrected for influences other than the saving actions. These so-called normalization factors can be **weather** (with effect on consumption), the **rebound** effect, performance gaps and changes in energy using **activities** per end-user, such as changes in the number of persons in the household. All normalization factors affect total savings through the unitary savings.

For the specific guidance at hand the normalization factors can concern weather, rebound effect and occupation rate. See table 1 in this link.

Here the normalization factors comprise weather and activity level, see in this link and in this publication and also here, or in this paper.

The possible **rebound effect and performance gaps** can be assumed to be **already included** in the "after" billing data. However, **it does not mean that billing analysis can distinguish these effects** in the results provided. Distinguishing the different effects and their magnitude requires additional data collection and analysis. For example, investigating rebound effect can be done through surveys to know if households have bought larger appliances or are now using more appliances. Investigating performance gaps can be done through direct measurements (see Specific Guidance 4).

When using the baseline Before/after, a correction must be applied on billing data for weather depending on the heating and cooling systems used by the households and the type of appliances under evaluation. The electricity consumption of most appliances is usually not significantly affected by changes in weather conditions from one year to the other. However, it can be sensitive to seasonality. Therefore, the billing data need to cover a long enough period (see section 5 about data issues). More importantly, in case the households use electric heating or cooling, then weather conditions will influence significantly the electricity consumption that therefore need to be corrected accordingly. More generally, billing might not be the most appropriate method in this case.

Also, a correction can be applied for changes in activity level. For example, in case of changes in the number of persons per household, or time spent at home (e.g. changes in employment). When these changes are significant, they can lead to outliers, which can make possible to identify them or to remove them from the sample used for the evaluation.

For the baseline "Target/Control group" all other normalization factors can be left aside as they are assumed to be the same for Target and Control group, given a sound sample set-up.

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 next section on "Calculating Gross and net savings).

Adjustment factors can concern the **Free rider** effect, the **Spill-over**/multiplier effect, **Additionality**, **Non-compliance** and **Double counting** (see table 1 here). Additionality and non-compliance are connected to unitary savings, while free riders and spill-over work through the number of actions.

In case of another policy focusing on the same saving actions as the mandatory energy labelling, the adjustment factor Double counting might be relevant. If the other policy is not covered in this guidance, double counting can only be accounted for at a higher level (i.e. policy mix or sectoral level).

See also link to note "Saving calculation methods and their application in the EPATEE Toolbox" and in the EMEEES project deliverable.

When dealing with a regulation, here mandatory energy labelling, it is difficult to define what the **free-rider effects** would be. In theory, free-riders would be households who would have bought the same appliance (with the same level of energy efficiency) in the absence of the energy labelling. However, in practice, this is very difficult to assess, as this is rarely possible to directly compare with a control group (except during trials or pilot phase). This can be done for example by making surveys or experiments where households would be in conditions to buy an appliance, with different settings (e.g. with/without energy labels, or with different types of energy labels). See for example (Bull, 2012; Fries et al., 2017; Mills and Schleich, 2010; Stadelmann and Schubert, 2018).

Similarly, the definition of **spill-over effects** of mandatory energy labelling is not straightforward, and these effects can be difficult to assess. Especially when there is no market or country that can really be comparable, and that the labelling scheme has been in place for many years. As the mandatory energy labelling is a regulation, it applies to the whole market for the types of appliances subject to this regulation. Spill-over effects could for example be:

1. manufacturers and retailers developing similar energy labelling for other types of appliances (not subject to the regulation);
2. manufacturers who will invest more on R&D, or more generally put more efforts to develop appliances with higher energy efficiency;
3. retailers who will include more appliances with higher energy efficiency on their catalogues or shelves, and more generally who will highlight more the energy efficient models in their marketing;
4. households who have been influenced by the energy labelling when buying an appliance and who will then pay more attention to the energy performance or energy consumption of other products they will buy;
5. similarly, households who will then pay more attention to the way they use the appliances.

Examples 1 to 3 are related to the evaluation of **market transformation** effects. Example 4 is related to the impact on households' **purchasing behaviours**. Example 5 is related to the impact on households' **energy-using behaviours**. For more details about possible approaches to investigate these effects, see for example (Vine et al., 2001; Vine et al., 2005).

Additionality is usually encompassing both, free-rider and spill-over effects. Additionality can be evaluated with the approaches described above for free-rider and spill-over effects, when the objective is to assess net energy savings. It can also refer to a pragmatic approach, with the objective to assess energy savings additional to a given scenario. In this later case, the evaluation will not use adjustment factors, but will define the baseline according to this scenario. The differences between the observed situation and this **baseline scenario** give the additional energy savings.

More generally, the evaluation approach should be adapted in case the energy labelling scheme has been **in place for many years**. In this case, it becomes indeed particularly difficult to assess what would have occurred in the absence of the mandatory energy labelling. For example, the free-rider effects in recent years can be partly due to spill-over effects from previous years. About this, see for example (Lane et al., 2007).

The adjustment factor **non-compliance** is also very important to consider, as mandatory energy labelling is a regulation. Non-compliance effects can have a major impact on the effectiveness of the labelling scheme. Examples of non-compliance effects can be:

1. when the actual energy performance of an appliance is lower than the one displayed on its energy label;
2. when manufacturers do not provide the data needed to display the mandatory information to be included in the energy label;
3. when retailers do not display the energy label at the point of purchase (e.g. stores, web marketplace).

Example 1 is equivalent to **performance gap** (see section 4.3 Normalization factors). The effect of this performance gap on the appliance's energy consumption will be reflected in the metered energy consumption, and is therefore directly included in the billing analysis. Comparing energy savings determined with billing analysis and deemed savings for the same type of appliance can help to identify possible performance gaps (and non-compliance). They can be further investigated as part of **market surveillance activities** (see for example Blomqvist and Fjordbak Larsen, 2015)).

Examples 2 and 3 can have an impact on **households' purchasing decisions**, as households will not have access to the energy labelling information. This can thus have an effect on the level of energy efficiency of the appliances bought, and thereby on the unitary energy savings. However, this cannot be detected by the billing analysis. This requires other types of investigations that can be part of market surveillance activities (e.g. checking information available at points of purchase, checking manufacturers' catalogue). As retailers and consumers are increasingly using web marketplaces, new technics based on "**web-crawling**" have been developed to facilitate such verifications (see for example Bennich et al., 2017).

In the case of the target/control baseline, most adjustment factors can be ignored as the comparison controls for the effects. However, as explained above in section 4.2 Calculation baselines, this case is rarely possible when evaluating mandatory energy labelling scheme.

4.5 Calculating Gross and Net energy savings

Gross savings represent the energy savings as perceived by the end-users. Net savings concern the savings that can be attributed to the policy measure (here the mandatory energy labelling).

When calculating the savings, a distinction must be made for most specific guides to the **unitary savings** and **number of actions**; see for example here.

The gross unitary savings can be calculated using the before/after baseline, taking into account the normalization factors. The unitary savings (i.e. per household) follow from the change in billed energy consumption before and after the implementation of saving actions (here installations of energy efficient appliances). The unitary savings will usually be defined per type of appliance and energy class. The number of actions will then need to be assessed with a complementary method, such as market survey or stock modelling (see section 2). When assessing gross savings, the assessment does not differentiate whether households' purchasing decisions were influenced or not by the energy labelling.

Net savings can be evaluated either directly (when using a control or comparison group, or a business-as-usual scenario as baseline) or from gross savings by applying further adjustment (or gross-to-net) factors. As discussed in section 4.2 Calculation baselines, the first case is rarely possible. Therefore, this guidance is focused on the case where the assessment of net savings will start from the same set of gross unitary savings per type of appliance and energy class. The difference will occur when considering the number of actions to calculate the total savings. The net number of actions follows from the gross number, corrected for the adjustment factors (possibly free rider effects and spill-over effects, see section 4.4 Adjustment factors).

Another approach to calculate net energy savings is to use econometric methods, when the types of data and related time series available enable to separate statistically the effects of the policy from other factors. See for example (Bjerregaard and Møller, 2019).

In case of applying the target/control group baseline, there is no need for correcting the (unitary) savings for normalization factors and for adjustment factors the same holds as for before/after.

Finally, total savings can be corrected for the Double counting effect (see section on adjustment factors), to be done at the level of the overall policy portfolios. For addressing double counting general discussions about gross and net savings, see for example, the IEA evaluation guidebook, or the EMEEES bottom-up evaluation methodology.

5 | INPUT AND OUTPUT

5.1 Main data requirements and data sources and collection techniques

Data requirements specified in the table below correspond to the calculation of energy savings, when using the baseline option **before/after.**

Calculation subject	Data requirements	Possible data sources and collection techniques
Energy consumption (excluding building related)	Billed energy consumption of participants (i.e. households who bought an efficient appliance during the period under evaluation) and fraction for building . The biiling date give of course an indication for total electricity , not for an individual appliance. So billing data give an estimate based on difference over a time period. Data points should be based as much as possible on meter readings (not on estimates) (see discussion below the table). (Note: billing analysis is usually made on a sample of households, as this is difficult in practice to collect billing data for all households who bought an appliance).	Bills from energy companies or from the households and (average) fraction for buildings per subsector from literature) In any case, a consent from the end-users (or utility waivers) is usually needed to comply with GDPR (General Data Protection Regulation) or equivalent regulation on data privacy and utility data.
Weather normalization	(*needed for households using electric heating or cooling*) 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 households (e.g. thanks to ZIP code). Weather normalization is in principal an easy task but especially during late spring and early autumn (normalization easily corrects too much or too less) it is not too easy to be made correctly. meaning this is one clear source of inaccuracy.	Data can usually be collected from the national weather services organisation.
Participants' data and data about the actions installed	Minimum data required is the type of appliance installed and its energy class. Other types of data needed depend on the regression model chosen, on the data needed to enable the extrapolation of the unitary savings to the number of actions, and possibly on complementary analysis planned. It is important to identify the data needs from the scheme outset, to enable the preparation of the data collection.	Data will usually need to be collected from a survey done specifically for the evaluation, possibly together with the contacts to collect the billing data.

Total number of actions	Number of appliances sold during the period of evaluation, per type of appliance and energy class.	These data may be available from national statistics. Or they will need to be collected from market surveys or bought from companies specialized in market data.
Fraction of the total sales that were influenced by the energy labelling	(*data needed when assessing net energy savings*) Fraction of households who bought an appliance, were aware of the energy labelling and took it into account in their purchasing decision. Effects on the purchasing decision (e.g. about the energy class of the appliance, its size, etc.)	Survey on sample of participants Econometric modelling (see for example Bjerregaard and Møller, 2019)

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. Using data from estimates can increase significantly the uncertainties. Therefore, it is recommended to use as much as possible data based on meter readings. With the roll-out of smart meters in most Member States, the availability of metered data should increase.

Smart meters can also provide data with very short timespans (every 15 minutes or less). Such data may make possible to disaggregate the load curve in a way that distinguish the main end-uses, including appliances. An example of use of high-frequency data can be found in (Chen et al., 2015).

However, these data are not used for billing. Collecting them usually require specific agreements with and consents from the households, as well as from the energy companies depending on the national regulation on smart meter data. An alternative can be to organize voluntary data “donation” by households. Such approach is usually combined with providing households with detailed energy advice as a feedback. We are not aware of such initiative for electricity consumption related to appliances. However, there are examples related to renovation works (CO2online; Milojkovic, 2016) and to smart thermostat (ecobee; Meier et al., 2019).

Similarly, web-crawling technics can be used as alternative way to collect market data about appliances (see for example Bennich et al., 2017; Enervee, 2014).

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.

As billing analysis uses metered data (i.e. data about final energy consumption), it first provides energy savings in final terms. It can also calculate savings in primary terms provided that billing analysis shows the savings per energy carrier and primary factors are available to convert the savings in final terms to savings in primary terms.

Most of the appliances covered by energy labelling use **electricity**. This means that the billing analysis is usually focused on electricity consumption. If the energy savings are calculated over the lifetime of the appliances (see below), then it is important to clarify if the primary factor used for electricity in the calculations takes into account the **likely changes in the electricity mix** over the period (e.g. due to the increase in the share of Renewable Energy Sources), or if the current primary factor is used for all years.

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 appliance can save electricity over its lifetime of about 10 to 15 years, LED over 10 to 20 years and more efficient computers up to 5 years.

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. Billing analysis is used to calculate unitary energy savings. It usually first provides annual energy savings, as the calculations are commonly made for twelve months. Then it can provide savings cumulated over the lifetime of the appliances, provided data are available or assumptions can be made about this lifetime.

However, **billing analysis does not provide alone the total savings**. Therefore, it cannot show the annual savings from the labelling scheme per type of appliances for a given year. For the same reason it **cannot provide directly cumulative savings according to the Article 7 in Energy Efficiency Directive**.

The complementary method(s) used to assess the number of actions will usually make possible to calculate total annual savings according to the years where appliances are bought or installed.

To evaluate total cumulative savings, data or assumptions about lifetimes per type of appliances are needed, as well as number of actions per type of appliances.

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.
- Ecodesign Impact Accounting – Status Report 2017. Prepared by VHK for the European Commission December 2017. (see annex A, pp.73-76).

For more details about lifetimes and savings persistence, see for example (Hoffman et al., 2015).

6 | ALTERNATIVE FOR CHOSEN METHOD

6.1 Alternatives for the chosen method

Billing analysis is commonly chosen when the evaluation objective is more specifically to assess or verify the energy savings actually achieved. **Billing analysis can particularly be used to evaluate savings in households who do not use electric heating or cooling**, and who bought appliance(s) representing a significant share of their electricity consumption (e.g. refrigerators). At the opposite, billing analysis would not be the most appropriate method in case consumption of appliances are mixed with consumption of heating or cooling devices. The respective consumption can be separated if monthly data of metered electricity consumption and external temperatures are available. However, this might add uncertainties to the result.

Another drawback of using billing analysis is the time lag between the installations of the actions and when the evaluation results can be available. Moreover, as explained in section 2, a complementary method is needed to assess total savings.

Deemed savings can be a more straightforward way to assess unitary energy savings, especially if based on differences in energy class. However, deemed savings hold more uncertainties than billing analysis: correction factors might indeed be needed to take into account possible performance gaps and rebound effect that are directly included in the billing analysis (see section 4.3 Normalization factors). For more details about deemed savings for household appliances, see Specific Guidance 14.

At the opposite, **direct measurements** are usually more costly to implement but provide more accurate unitary energy savings, including for appliances representing a small share of the households' electricity consumption. Direct measurements also make possible to distinguish performance gaps (difference in savings due to technical aspects) and rebound effect (difference in savings due to behavioural aspects). Which billing analysis cannot (except if additional data are collected, which would then be similar to an approach based on direct measurements) see Specific Guidance 4.

Then **complementary methods** can be used to assess the number of actions: stock modelling (see Specific Guidance 5), diffusion indicators (see Specific Guidance 7), market surveys, etc. (see section 2).

The table below presents the pros and cons of billing analysis (covered in this guide) and for commonly used alternative methods for the same combination of policy measure and sector.

Type of method	Pros	Cons
Billing analysis (Method covered in this advanced guide)	Billing data showing actual changes in energy consumption over time. Performance gaps and rebound effect directly taken into account (This feature of billing data will depend on country)	Total energy consumption to be corrected for building-related energy use in case of households using electric heating or cooling, thus will difficult to get reliable results and will result in huge uncertainties. Difficulties to get statistically significant results for appliances representing a low share of households' electricity consumption. Need for large enough samples to get statistically significant results. Complementary methods needed to assess the number of actions or "participants" (i.e. to extrapolate savings from the sample to the whole households subject to the labelling scheme).
Deemed savings (see Specific Guidance 14)	• Straightforward to apply (once a set of deemed savings has been defined) • Can be used for ex-ante evaluation and with small, or even no, time lag for ex-post evaluation • Savings per type of appliances and energy class	• Need to define a set of deemed savings to cover the various types of appliances and energy class • Uncertainties due to the assumptions used • Correction factors needed to take into account possible performance gaps and rebound effect • Complementary method required to assess the number of actions (cf. market shares at least disaggregated per energy class)
Direct Measurement (see Specific Guidance 4)	• High accuracy for the data on energy consumption • Data that reflects actual use of the appliances (if on-site measurements; not for laboratory tests) • Possibility to identify non-compliance (especially for laboratory tests)	• Expensive and time-consuming • Difficult to obtain data that are representative of the whole market or stock of appliances. • Complementary methods needed to assess the number of actions (and market shares at least disaggregated per energy class) • Analysis only after physical implementation of savings action (if on-site measurements)

Diffusion indicators (see Specific Guidance 7)	**Detailed analysis about the number of actions, possibly enabling an evaluation of market transformation effects**	**Complementary method needed to estimate the energy consumption per type of appliances**
Stock modelling (see Specific Guidance 5)	Detailed analysis about the number of actions, possibly enabling an evaluation of market transformation effects In case available, possibility to calibrate the model with national statistics about energy consumption of household appliances	Complementary method needed to estimate the energy consumption per type of appliances
Consumer surveys	Provide a better understanding of the effects of the labelling scheme, and thereby useful feedback to improve the scheme	Difficulty to obtain representative samples and quantitative results that can be used to calculate energy savings (unless large and repeated surveys can be done)

See also section 5 about alternative methods for data collection.

7 | ADDITIONAL EVALUATION RESULTS

7.1 Calculating avoided CO_2 emissions

Avoided CO_2 emissions can be evaluated from the energy savings by applying emission factors. Four key aspects are to be taken into account when choosing the emission factor(s):

1) Emission factors vary according to the **energy type**, so the data about energy savings need to be available per energy type.
2) Emission factors for a given type of energy **can vary over time** (especially for **electricity**).
3) Emission factors can take into account:
    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.

When the savings concern one fuel only, e.g. gas boilers, the reduction in CO_2-emissions can be calculated from the savings with an emission factor for the fuel at stake. The reduction in CO_2-emissions can only be calculated when savings are calculated per relevant energy carrier and a specific emission factor is available for each energy carrier.

Most of the appliances covered by mandatory energy labelling use electricity. It is important to document what emission factor is used for electricity, as there can be different ways to calculate this emission factor. The conversion of electricity savings into CO_2 savings is indeed a special case, depending on the national mix for electricity production. Examples of different approaches are:

- **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** (also called **marginal emission factors**): 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).

The choice between these two options depend on the national electricity mix (cf. emission factor varying significantly with time of production or not) and the type of appliances examined by the evaluation.

Moreover, when cumulating savings over time, 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). A pragmatic approach to keep the calculations relatively simple should be the leading principle.

The avoided emission of **other greenhouse gasses** due to energy savings are not taken into account here, as these emissions (and more specifically their reductions) are generally negligible compared to CO_2 when dealing with household appliances, except in special cases (e.g. if it can be shown that the appliances promoted by the labelling scheme can lead to reductions or increases in leakages of refrigerant fluids).

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 **manufactures' or retailers'** point of view (e.g. taking into account extra R&D investments, changes in turnover or mark-up)
- Cost-effectiveness for society at large (e.g. social net present value)
- Cost-effectiveness from the point of view of the **public authority** (e.g. comparing different types of policy measures)

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

The calculation of cost-effectiveness for end-users demands, next to the savings, data on investments made, subsidies on investments (if any), interest rates (if credits were used to buy the appliances), lifetimes of the saving actions, energy prices (including taxes) per type of end-user and discount factors (possibly differentiating different types of households).

For cost-effectiveness from a societal viewpoint no account is taken of subsidies and taxes, energy prices concern world market price, and a lower value of the discount factors is valid (e.g. using the reference discount factor defined for public investments).

For cost-effectiveness from a government perspective, the analysis should take into account the administration costs of the energy labelling scheme (e.g. defining and updating the energy classes, market surveillance activities). If the effects of the labelling scheme are expected to be large, it can also be relevant to consider changes in tax revenues due to the scheme (see table below). There can also be other public costs to take into account if accompanying measures are implemented (e.g. subsidies, costs of information campaign).

The difference of perspectives between society and public authorities is summarized in the table below.

Point of view	Costs	Benefits
Public authorities	• Administration costs • Costs of accompanying measures (e.g. subsidies, information campaign) (when relevant) • Losses in tax revenues related to energy taxes (due to additional energy savings)	• Additional or net energy savings
	• Changes in tax revenues related to VAT on appliances	
Society	• Administration costs for the public authorities	• Additional or net energy savings
	• Marginal investments in new appliances (compared to the business-as-usual scenario)	

NOTE: the table above does not deal with **non-energy impacts**. Depending on the context and objectives of the energy labelling scheme, non-energy benefits can be important to include in the cost-benefit analysis (e.g. reductions in GHG emissions, reductions in energy poverty and bills arrears).

When assessing the cost-effectiveness of a labelling scheme, several challenges can arise:

- The scope of administration costs (and possibly of accompanying measures) can be difficult to define, and corresponding data can be difficult to collect or establish (e.g. part of the administration cost is related to time spent by civil servants; market surveillance activities can be common to other policy measures).
- Prices of appliances can be difficult to monitor in a systematic way, and it can be challenging to define a business-as-usual scenario including assumptions about trends in costs of appliances.

Usually, cost-benefit analyses of energy efficiency policies assume that the more energy efficient options will be more expensive than the less energy efficient ones in terms of investment costs. Therefore, they include as costs the additional investment costs. When dealing with appliances, this is not always the case. Sometimes, more efficient appliances can have the same or even smaller investment costs than less efficient appliances (for example, when the more efficient appliances become dominant on the market). That is why in the table above, the parameters related to the prices of appliances can either be costs or benefits.

When using **discount factors** (e.g. when the calculated indicator is Net Present Values), it is important to document the values used for discount factors, and if possible to make a sensitivity analysis (testing several values or ranges of discount factors) as this can significantly affect the results.

Likewise, the calculations of cost-effectiveness indicators will usually require to consider scenarios of **energy prices** over given periods. The assumptions about trends in energy prices should be documented. Whenever possible, it is recommended to make a sensitivity analysis (testing several scenarios of energy prices).

In practice, as it is difficult to separate the effects of energy labelling and MEPS (minimum energy performance standards), the cost-effectiveness analysis will be made for the whole policy package (energy labelling + MEPS). In that case, the business-as-usual scenario should be adapted accordingly and the administration costs of the MEPS should be taken into account together with the administration costs of the labelling scheme.

7.3 Calculating other co-benefits

Possible co-benefits from saving energy concern:

- Extra employment
- Reduction of energy poverty
- Other emission reductions (NO_x, SO_2, fine particles, etc.)
- Better indoor climate
- Reduced dependency on (insecure) energy import

In practice, co-benefits and non-energy impacts have rarely been subject to evaluation when dealing with energy labelling schemes. And like for other energy efficiency policies, non-energy impacts have been more often evaluated as part of impact assessments (ex-ante evaluations) than in ex-post evaluations. The following examples were found in the review of ex-post evaluations of standards and labelling programmes done by Cory Smith et al. (2016).

Type of non-energy impacts	**Why it can be relevant (and for whom)**	**Case(s) mentioned by Cory Smith et al. (2016)**
R&D and technology developments	Investments in R&D	Europe (ex-ante)
Industrial competitiveness	Important criteria for policy makers, and to engage private stakeholders	Europe (ex-ante)
Job creation	Important criteria for policy makers	Europe (ex-ante)
Environmental impacts (water use, noise, air pollution)		Europe (ex-ante) US (ex-ante and ex-post evaluation of the US standards for residential and commercial appliances).

Providing guidance for the evaluation of non-energy impacts goes beyond the scope of this guidance. Examples of such evaluations can be found in the references mentioned above.

8 | CONCRETE EXAMPLES

A comprehensive study that among others treats household appliances has been performed by Joanne Wade and Nick Eyre (2015) for the UK Energy Research Center. See more specifically the section 4.3 about appliance market transformation activities.

The EPATEE case studies do not include an example to use billing analysis to assess energy savings from household appliances.

We could not find in the literature either any practical example of using billing analysis to assess energy savings from household appliances. However, this does not mean that such example does not exist. As our time for literature review was limited.

9 | FURTHER READING

General guidance on evaluations

- 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.
- 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.
- Eichhammer et al., 2008. Distinction of energy efficiency improvement measures by type of appropriate evaluation method. Final Report on Evaluation and Monitoring for the EU Directive on Energy End-Use Efficiency and Energy Services. EMEEES Project report, https://ec.europa.eu/energy/intelligent/projects/sites/iee-projects/files/projects/documents/emeees_project_report_en.pdf
- 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.
- Khawaja, S., Rushton, J., & Keeling, J. (2017). Chapter 11: Sample Design Cross-Cutting Protocol. The Uniform Methods Project: Determining Energy Efficiency Savings for Specific Measures. Prepared for NREL (National Renewable Energy Laboratory), September 2017.
- Vreuls, H. et al., 2005. Evaluating Energy Efficiency Policy measures & DSM programmes. Volume II Country reports and case examples used for the evaluation guide book. Prepared under the IEA Implementing Agreement on Demand-Side Management Technologies and Programmes.
- 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 mandatory energy labelling

- Ecofys, Waide SE, SoWatt, Öko-Institut, SEVEn & University of Coimbra (2014). Evaluation of the Energy Labelling Directive and specific aspects of the Ecodesign Directive. Final technical report for the European Commission, June 2014.
- European Commission (2019). Clearer and simple energy labels will help consumers save money and contribute to the Energy Union's objectives.
- Hirayama, S., Nakagami, H., Murakoshi, C., Nakamura, M., Mizutani, S. & Okuda, S. (2008). International Comparison of Energy Efficiency Standard and Labels: Development Process and Implementation Phase. Proceedings of the 2008 ACEEE Summer Study on Energy Efficiency in Buildings, 8-110:8-122.
- IEA 4E (2016). Achievements of appliance energy efficiency standards and labelling programs: A global assessment in 2016. Report of the IEA Technology Collaboration Programme on Energy Efficient End-Use Equipment (IEA 4E), August 2016.
- Kemna, R. (2014). Ecodesign impact accounting – Part 1. Report by VHK for the European Commission. 30 May 2014.

- Tholen, L., Adisorn, T. & Götz, T. (2017). It's nothing to write about? How EU-Member States address appliances in their National Energy Efficiency Action Plans. Proceedings of EEDAL (Energy Efficiency in Domestic Appliances and Lighting) 2017, 171-182.
- Waide, P., Watson, R., Attali, S., Eide, A., Krivosik, J. & Schiellerup, P. (2013). The new energy label: assessing consumer comprehension and effectiveness as a market transformation tool. Proceedings of the ECEEE 2013 Summer Study.
- Wiel, S., & McMahon, J.E. (2005). Energy-Efficiency Labels and Standards: A Guidebook for Appliances, Equipment, and Lighting (No. LBNL-45387-2nd-Edition; LBNL-45387). Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA (US).

About methodologies to evaluate impacts from standards and labelling schemes for households appliances:

- Corry Smith, J., Yu, Y., Carreño, A.M. (2016). Ex-Post Impact Evaluations of Appliance Standards and Labelling Programmes: A Global Review of Best Practices and Lessons Learned. Proceedings of IEPPEC 2016.
- De Melo, C. A., & de Martino Jannuzzi, G. (2010). Energy efficiency standards for refrigerators in Brazil: A methodology for impact evaluation. *Energy Policy*, 38(11), 6545-6550.
- Lane, K., Harrington, L. & Ryan, P. (2007). Evaluating the impact of energy labelling and MEPS – a retrospective look at the case of refrigerators in the UK and Australia. Proceedings of the ECEEE 2007 Summer Study, paper 4-292, 743-750.
- Vine, E., Choi, J.Y., du Pont, P. & Waide, P. (2005). Chapter 9: Evaluating the impact of energy – efficiency labelling and standard-setting programs. In: Wiel, S., & McMahon, J. E. (2005). Energy-Efficiency Labels and Standards: A Guidebook for Appliances, Equipment, and Lighting (No. LBNL-45387-2nd-Edition; LBNL-45387).
- Vine, E., du Pont, P., & Waide, P. (2001). Evaluating the impact of appliance efficiency labelling programs and standards: process, impact, and market transformation evaluations. Energy, 26(11), 1041-1059.

Studies about the effects of energy labelling on purchasing decisions

- Bull, J. (2012). Loads of green washing—can behavioural economics increase willingness-to-pay for efficient washing machines in the UK? *Energy policy*, 50, 242-252.
- Fries, B., Braungardt, S., Kreuzer, M. & Wohlfahrt, K. (2017). EU product policy and consumer purchase decisions – empirical evidence from eight EU member states. Proceedings of the ECEEE 2017 Summer Study, 1593-1602.
- Mills, B., & Schleich, J. (2010). What's driving energy efficient appliance label awareness and purchase propensity? *Energy Policy*, 38(2), 814-825.
- Stadelmann, M., & Schubert, R. (2018). How do different designs of energy labels influence purchases of household appliances? A field study in Switzerland. *Ecological economics*, 144, 112–123.

Examples about distinguishing effects from energy labelling and other policy measures:

- Bjerregaard, C., & Møller, N. F. (2019). The impact of EU's energy labelling policy: An econometric analysis of increased transparency in the market for cold appliances in Denmark. Energy Policy, 128, 891-899.
- Sulyma, I. & Tiedemann, K.H. (2015). Evaluating the Combined Impact of Financial incentives, Tax Incentives, Information and Minimum Energy Performance Standards in Transforming a Residential Appliance Market. Proceedings of EEDAL (Energy Efficiency in Domestic Appliances and Lighting) 2015, 1186-1196.

About new data collection techniques related to data on appliances

- Bennich, P., Stengard, L. , Christensen, S.F., Hartikainen, T., Mogensen, K. & Fjordbak Larsen, T. (2017). Using webcrawler techniques for improved market surveillance – new possibilities for compliance and energy policy. Proceedings of the ECEEE 2017 Summer Study, 1633-1642.
- Enervee (2014). Recent and Historical Product Energy Efficiency and Life-cycle Cost Improvement in Swedish Appliance Markets. January 2014.

About alternative sources and types of analysis related to data on energy consumption

- Chen, V. L., Delmas, M. A., Kaiser, W. J., & Locke, S. L. (2015). What can we learn from high-frequency appliance-level energy metering? Results from a field experiment. *Energy Policy*, 77, 164-175.
- Meier, A., Ueno, T., Pritoni, M., Rainer, L., Daken, A. & Baldewicz, D. (2019). What can connected thermostats tell us about American heating and cooling habits? Proceedings of the ECEEE 2019 Summer Study, paper 4-042-19, 649-658.
- Milojkovic, F. (2016). Refurbishment Effectiveness Assessment Based On Consumption Data From iESA Accounts. Proceedings of IEPPEC 2016.

About issues related to compliance

- Blomqvist, L. & Fjordbak Larsen, T. (2015). Effect project – estimating the benefits of Nordic market surveillance cooperation on ecodesign and energy labelling. Proceedings of the ECEEE 2015 Summer Study, paper 8-266-15, 1837-1848.
- Lock, T., Muir, S., Faninger, T., Tinetti, B. & Krivosik, J. (2015). What's on? Compliance of Televisions with Energy Labelling and Ecodesign Regulations. Proceedings of EEDAL (Energy Efficiency in Domestic Appliances and Lighting) 2015, 247-258.
- NEA, 2015. Results of Verification Testing of Registrable Goods Under the Mandatory Energy Labelling Scheme. Report of Singapore's National Environment Agency, January 2015.

Acknowledgments & Disclaimer

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

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of the following information. The views expressed in this publication are the sole responsibility of the author and do not necessarily reflect the views of the European Commission.