

WHAT CAN WE LEARN FROM SHARING EXPERIENCE ABOUT EVALUATION PRACTICES?

Gregor Thenius

Österreichische Energieagentur, Mariahilferstraße 136, +43 1 586 1524,
gregor.thenius@energyagency.at, www.energyagency.at

Kurzfassung:

Background

Many resources are available about state-of-the-art or best examples of evaluations of energy efficiency policies and evaluation guidelines, like in the IEPPEC conference and IEPPEC proceedings. However what do we know about daily evaluation practices? Is it always easy to find examples of evaluations about a given country or type of policy instrument? Are evaluations used to improve policies? And finally, how evaluation practices could be improved? This is the kind of issues that the European project EPATEE (www.epatee.eu) aims at tackling, focusing on impact evaluations.

Getting feedback from stakeholders

About 60 **stakeholders** from 13 European countries were **interviewed or surveyed** to better know their priorities about evaluation issues, how they would define the level of evaluation practices in their country and what barriers would impede effective evaluation practices.

Stock taking on existing evaluation practices

In parallel, **references were collected** and coded to build a knowledge base gathering already more than 170 evaluation reports, papers or guidebooks. In a further step **case studies** are analysing around 30 evaluations to provide concrete and detailed experience feedback about why evaluation is used, how it is performed and what difficulties are encountered.

The objective of the project is not to provide an exhaustive or representative picture of the evaluation practices in Europe, but to gather and develop materials that can be used as a basis for **experience sharing activities**, as well as to develop an online tool box that will make these resources available in a user-friendly way. The key assumption of the project is that concrete examples and guidance can help overcome barriers that currently limit the use of evaluation. Experience feedback indeed shows that evaluation can be a very effective tool to improve policies, thereby achieving more energy savings at lower costs.

This paper presents the results of the first phase of the project, focusing on the main conclusions from the stakeholders' survey, the knowledge base and the case studies. Feedbacks gathered remind usual no-brainers (e.g., anticipating data collection). It also shows that evaluation is not only a technical issue, but that organisational issues (e.g., cooperation between institutions) are critical as well. Stakeholders expressed their interest in evaluation of cost-effectiveness of policies, and effects other than energy savings.

Keywords: Impact evaluation, Energy efficiency policies, Energy savings

1 Background

Directives on energy efficiency such as the EED or the EPBD trigger a great variety of policies throughout EU Member States. The effort put into developing and implementing these policies is well documented for example in existing National Energy Efficiency Action Plans. Soon, emphasis will be put on finding out how effective they were to create new policies that allow meeting current and future energy savings targets. Unfortunately, several barriers limit policy evaluation. This results in a lack of quantitative data, and impedes evidence-based analysis required to distinguish effective from ineffective policies.

This problem can be tackled by raising the capacity of policymakers and implementers to assist Member States to fulfil their obligation under energy related Directives. The project EPATEE provides them both with tools and with practical knowledge to make effective impact evaluation an integral part of the policy cycle. EPATEE makes use of existing evaluation experiences in a range of instruments, such as energy efficiency obligation schemes, regulations, financial incentives and voluntary agreements. Lessons learnt from other EU initiatives and good practices in how to successfully evaluate the impact and cost-effectiveness of such energy efficiency policies will provide the basis for the development of guidelines and good practice evaluation tools.

This report shows the project's results so far with respect to the identification of stakeholder needs and existing literature and practice in energy efficiency policy impact evaluation.

2 Stakeholder needs and experiences

Knowing stakeholders needs and experiences is key to providing them with tailor-made support to overcome obstacles for an effective policy evaluation. The results presented in this chapter are based on:

- A series of 26 interviews with a group of key stakeholders identified in EU-Member States to get a qualitative feedback and identify basic needs and priorities on impact evaluation;
- An online survey among a larger group of stakeholders, aimed at collecting more quantitative feedback, better understanding the needs of the stakeholders and ranking the priorities identified in the interviews.

The importance of impact evaluation

All the stakeholders agree that ex-ante and ex-post evaluation contribute to improve energy policies (both for the design of new ones or the revision of existing ones). The importance of ex-ante and ex-post evaluation is seen equal. Some stakeholders suggest that the importance of evaluation becomes obvious if one thinks about what would happen if evaluation is not done: effects of the policies would not be known and the efficiency of the use of public budget could not be assessed.

Many stakeholders have examples of evaluations which led to an improvement of policies. Nevertheless, the type of analysis and the extent of the monitored and studied effects vary depending on

- The type and size of the policy;

- The priorities of policy makers;
- Other conditions related to special framework conditions in countries.

Time and other resource constraints often impede the possibility to cover all evaluation needs and to ensure a complete and reliable analysis. This explains why some of the stakeholders suggest that evaluation should be made mandatory for all major policies and resources should be allocated to evaluation already in the design phase of policies.

Another related issue is the need to explain the distinction between monitoring and verification (M&V) and evaluation, and more specifically to show what added value evaluation brings compared to M&V.

How much standardisation?

Some stakeholders would like to have standardized tools (guidelines, web based information systems, etc.) in order to be able to carry out evaluation more easily (both ex-ante and ex-post), but also to make it easier to compare evaluation results between policies and countries. Such an approach can be useful for decision makers who want to optimize policy portfolios or to prioritize policy efforts in terms of financial resources. On the other hand, there are strong reservations against too much standardization. Evaluation often needs to be tailored to the objectives and context of the individual policy and evaluation practices that prove effective in one country can not necessarily be transferred to other countries. One potential compromise with view to standardisation is that at least standard guidelines are used to report evaluation results in order to make them more transparent and comparable.

Persistent data issues

One of the persistent barriers to meaningful evaluation is related to the availability of the right data and information at the right moment. However this lack does not only include missing data related to the energy efficiency actions themselves but also to qualitative information about the policy background. Neglecting the policy background and the framework a policy is embedded in by external evaluators may lead to misinterpretations of the results.

Solutions to overcome these data and information issues include:

- The use of ICT tools to collect and process quantitative data;
- Planning evaluation early enough to ensure the feasibility of ex-post evaluations. Even if this basic rule is well-know, it is not necessarily put into practice according to stakeholders.

Effects like energy savings, the number of individual actions, the use of public resources, CO₂ emission reductions and the evaluation of results against targets are more often evaluated than effects such as market effects, employment, or the qualification of market operators. Linking these results with barriers identified one reason for this situation could lie in the lack of relevant data as well as in the lack of well-established methodologies to evaluate non-energy benefits of policies. This suggests that work needs to be done to introduce the evaluation of such equally important issues in the evaluation cycle.

Net effects

Another challenge is the evaluation of net effects of policies, i.e. how effects of a policy can be separated from other effects (other policies or external factors). This has proven to be challenging particularly for energy efficiency policies.

Organisational barriers

From the feedback gathered in the interviews and the survey it can be concluded that organisational barriers in impact evaluation are often seen as more severe than technical issues. The two main barriers stakeholders are often confronted with in impact evaluation are:

- Insufficient financial resources for evaluation;
- A lack of interest from policy makers and public managers.

Both barriers eventually lead to a situation where learning about policies and their effects by evaluating them is being made difficult or even impossible. Based on this insight the project EPATEE will focus in its activities on showing to policy makers the possible value-added of evaluations.

Support needed

According to stakeholders many aspects of the evaluation process require support and tools, such as:

- Concrete good practice examples of evaluations;
- Links to current evaluation reports;
- Monitoring and verification methods;
- Indicators to verify the cost-effectiveness of policies;
- Data and approaches needed to analyse non-energy effects of energy efficiency policies (e.g. enterprise competitiveness, fuel poverty, environmental and social benefits, rebound and free riders' effects);
- Exploring opportunities of harmonization or of standardised procedures to evaluate policies in a comparable way keeping in mind to take into account different framework conditions;
- Ensuring that evaluation is an independent process and is not biased, so stakeholders can trust evaluation results;
- Ensuring cost-efficiency of evaluation by balancing resources employed with results expected.

The EPATEE project will implement a set of tools to tackle some of these needs and organise workshops and webinars to facilitate the sharing of information and experiences among policy makers, evaluators, experts and the other stakeholders involved in policy evaluation.

3 Existing literature and experiences

3.1 Building up a knowledge base

Although a lot of information is already available in the field of policy evaluation, it is usually not easily accessible. To close this gap, the project EPATEE produces a synthesis about the knowledge and experience available and a summary about the main issues and gaps to tackle for the development of evaluation practices. Collecting and structuring available knowledge in a user friendly way helps to overcome the barrier that the rich material available may be difficult to use by the stakeholders, due to lack of time, diversity and dispersion of the information sources and complexity of the topic.

Thus in parallel to identifying stakeholders views and needs with regard to evaluation the project EPATEE looked at existing evaluation practices by analysing reports on concrete evaluations as well as meta-studies on this topic. In order to analyse all documents in a harmonised way, a framework for extracting the relevant information was designed. The criteria analysed include:

- Type of policy instrument
- Sector addressed by the policy
- Evaluation type (e.g. ex-ante, ex-post)
- Objective of evaluation (impact or process evaluation)
- Data Collection
- Calculation method
- Baseline (counterfactual)
- Presentation of data on energy savings (gross vs net)
- Normalisation factors
- Adjustments of energy savings
- Is cost data available? Is the scope of the cost data explicit?
- Uncertainty analysis done
- Further impacts of measure/policy

To make this information available a user-oriented knowledge base will be built, including a mapping of the resources available (“where to find what”) thereby ensuring an easy access to these resources. An online tool making this existing knowledge available more easily to stakeholders is in preparation and will be ready to use in the first half of the year 2018.

3.2 Case studies on evaluation practices

“Our experience with the ex-post impact evaluation is that it’s really worth the effort”. This statement from one of the evaluation customers interviewed within the project EPATEE indicates the potential for well-executed evaluations and their benefits for the evaluated policy. But how to transfer these insights to other policy makers?

The experience sharing about evaluation practices is limited due to the lack of time for stakeholders to disseminate or document their evaluation works, and due to the many languages in the European countries. Therefore, the actual evaluation practices of the stakeholders are not well known, and most of the evaluation results are disseminated without

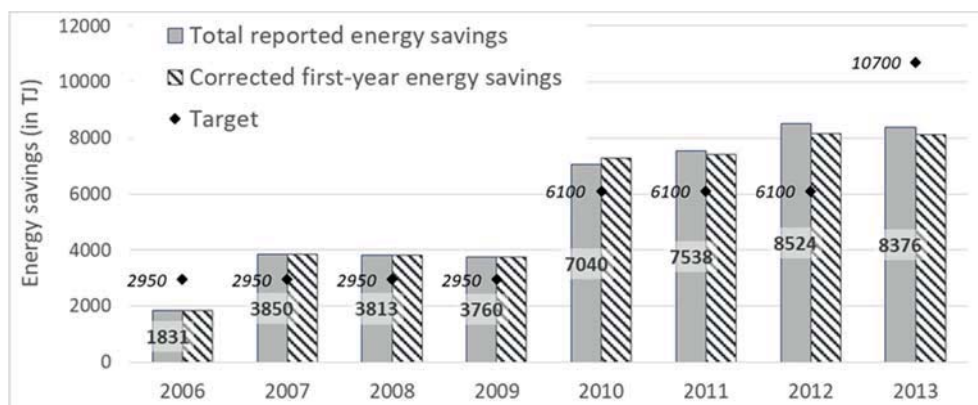
explanations about their evaluation methods. This creates limitations for a correct understanding and use of evaluation results. This is a barrier to know what policies are effective and efficient, and may create doubts in the reliability of their results. To tackle this issue the project EPATEE produces a set of case studies on recent ex-post evaluations, using a common methodology and template (<https://epatee.eu/case-studies>). The cases are selected based on the expertise of the partners, the priorities expressed by the stakeholders and taking into account the issues identified in creating the knowledge base (see chapter 3.1). These case studies on concrete evaluation examples will be complemented by topical case studies on horizontal issues of interest (e.g. gross vs net savings, the linkages between monitoring and verification and evaluation). The following two chapters show results from case studies on Danish and Irish evaluations.

3.2.1 The Danish energy efficiency obligation scheme

The evaluation of the Danish energy efficiency obligation scheme (EEOS) is an example of a long-lasting scheme that underwent a number of evaluations and subsequent changes in its design.

The objective of the scheme is to promote cost-effective energy savings that would otherwise not have been realized. Energy distributors are required to achieve yearly energy savings targets, and must report each year their achievements to the Danish Energy Agency that undertakes random controls of the reported measures. Energy distributors may provide advice and information about energy savings, implement energy savings projects on their own grid system or via meters, establish agreements with contractors that will implement programmes towards end-users or provide subsidies to end-users through direct contracts. Eligibility criteria for energy efficiency measures include minimum energy performance requirements and rules about additionality.

In the following table **reported energy savings** are first-year energy savings including conversion factors, and from 2010 reduction and prioritisation factors. Conversion factors are applied in cases of a substitution between energy sources; reduction factors are based on additionality assessments done in previous ex-post evaluations; prioritisation factors are defined to favour actions with longer lifetime, having impacts on primary energy consumption and in terms of avoided CO₂ emissions. **Corrected first-year energy savings** are first-year energy savings including conversion and reduction factors but not prioritisation factors.



Source: https://epatee.eu/sites/default/files/epatee_case_study_denmark_eeo_scheme_vfinal2.pdf

The use of reduction factors already shows that prior evaluations are taken into account when defining rules for future obligation periods. The sources of uncertainties about energy savings are:

- errors in the calculations and reporting of the energy savings (tackled by random checks);
- uncertainties related to the use of engineering calculations or deemed savings (e.g., differences between estimated and observed energy consumption);
- uncertainties related to the reduction factors.

Ex-post evaluations of the scheme were performed in 2008, 2012 and 2015. The main objectives of the ex-post evaluations were to investigate whether the rules of the scheme were appropriate, the level of satisfaction of the stakeholders (obligated parties, end-users, etc.), the costs induced by the scheme and its overall cost-effectiveness. The following parts of this chapter focus on the evaluation of the additionality of energy savings.

Assessing additionality is key to ensure that the scheme delivers a net benefit to end customers. Additionality was defined in the evaluations as follows: *“energy savings are additional if the energy savings actions had not been implemented (today or for example within the next few years) in the absence of the obligation scheme”*.

In the evaluations done in 2008 and 2012, additionality was assessed by a survey of a sample of participants (companies and households) asking them:

- to what extent were you thinking about realising the energy saving project, before you were in contact with the obligated party?
- with what probability the project would have been realised within the next year (or within three years) without the help from the utility?
- In 2012, these questions were complemented to check the consistency of the answers: How critical to the implementation of the project was the subsidy you received?

In 2012, changes in energy consumption of a control group and participants group were compared over 24 months, showing that the net effect for the participants group would be about 56% of the energy savings reported. Such approaches always raise concern about the representativeness of the results because of the sample size but also about the reliability due to a possible bias in the answers to this type of hypothetical question. Using a few questions may be a cost-effective approach, but can be questioned for validity reasons. The evaluators recommended for the next evaluations to add qualitative questions and a plausibility check with less subjective methods to assess the baseline such as market data.

The evaluation of the year 2015 included the following new features compared to previous evaluations:

- A web-based survey with a control group (households and “non-households”). The reason for taking a web-based approach was that it provides more reliable answers compared to the previous evaluations by phone survey, as respondents get the opportunity to thoroughly consider the answers they submit;

- An econometric analysis of variables related to end-users and energy savings projects, to investigate what variables (within each end-user group) may influence the additionality rates;
- A top-down regression analysis per end-use sector, based on long time series of main macro-economic variables affecting energy consumption and costs incurred by the energy distributors for the scheme, to assess the overall net effects (assuming that this method enables to capture directly or indirectly the additionality of energy savings projects, rebound effect and spill-over effect);
- A survey of contractors (installers and engineering consultancies) assessing whether there have been spill-over effects in the supply chain. The evaluation concluded that this approach does not provide an accurate estimation of spill-over effects, but brought some insights.

The focus on the evaluation of the Danish EEOS on additionality is remarkable as it tackles a central question for energy efficiency policies and their impacts.

3.2.2 The Irish Better Energy Homes Scheme

The Better Energy Homes (BEH) scheme aims at improving the energy efficiency of dwellings, reducing heating bills, CO₂ emissions and air pollutant emissions. It provides direct Government grants representing about 30% of the total investment costs to homeowners and landlords renting dwellings to upgrade their dwellings with energy efficiency actions that must be installed by qualified professionals. Eligible action types are ceiling/attic insulation, inner and outer wall insulation, heating controls, high efficiency boiler upgrades and solar heating systems with minimum performance criteria and technical requirements for each action type. Actions receiving a BEH grant can also get support from an energy company within the Irish Energy Efficiency Obligation scheme.

The ex-post evaluation was used to prepare a public communication showing to households the actual energy savings achieved thanks to the programme. The ex-post evaluation of actions undertaken in 2009 was focused on the two following questions:

- How much energy savings were realised by people who had made energy efficient home improvements under the BEH scheme?
- How close were the actual energy savings realised to the technical savings potential forecast when the BEH scheme was set up?

This ex-post evaluation was based on the analysis of metered gas consumption (at least 2 reads per year) of 2 samples:

- a participants group: 210 homeowners who invested in actions with a BEH grant, and who made an energy assessment of their dwellings before and after the works;
- a control group: 153,928 households with similar dwellings in terms of type who did not participate in the BEH scheme.

In addition to the metered data, the Building Energy Rating database and other national statistics were used to find data about dwellings' characteristics in particular for matching "control" households with participants. A survey of the participants collected more detailed data related to heating demand and behaviours for further analyses.

The statistical approach used to evaluate the energy savings was based on a difference-in-difference method (quasi-experimental approach), comparing pre- (2008) and post-intervention (2010) heating consumption for both groups. The use of a control group was thus meant to ensure that the energy savings evaluated were related to the improvements carried out by the homeowners, and not to other factors that might affect all gas users (e.g. fluctuations in usage relating to price or to extreme weather conditions). Using the difference-in-difference method implies that the baseline is the change in energy consumption observed for the control group between 2008 and 2010.

The reasons for choosing the difference-in-difference method versus other possible options were:

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

In parallel, the billing data analysis was compared with simplified engineering calculations for energy savings. The ex-post evaluation showed final energy savings of about 21% for the participants on average compared to the control group. Compared to the ex-ante estimate of predicted average final energy savings per dwelling the ex-post results were about 36% lower. This may be due to the effects of behavioural changes (direct and indirect rebound effects), poor initial estimates of achievable savings (for example due to ex-ante assumptions) and poorly performing equipment and potential inefficiencies in the systems installed. The evaluators thus pointed out greater comfort among the co-benefits of the energy efficiency improvements. Some of these lifestyle improvements can explain part of the gap between the ex-ante engineering estimates and the ex-post billing analysis, but not all.

The Irish case study clearly shows the value added of an ex-post impact evaluation which led to a more realistic assessment of energy savings realised within the scheme.