EPATEE

Evaluation into Practice: Lessons learnt from 23 evaluations of energy efficiency policies

Volume II: Background report

Project Coordinator: Austrian Energy Agency – AEA

Work Package 3 Leader Organization: IEECP (Institute for a European Energy & Climate Policy)

August 2018

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

Authors

Jean-Sébastien Broc, IEECP

Vlasis Oikonomou, IEECP

Mia Dragovic, IEECP

With contributions by (authors of the case studies used to prepare this report):

Catherine Guermont, Laurence Ould-Ferhat, Marie-Laure Nauleau, Isabelle Sannié, Gregory Chedin, Nathalie Martinez and Gérald Lalevée (ADEME), Gregor Thenius (AEA), Paul van den Oosterkamp, Jamilja van der Meulen, Marijke Menkveld, Karina Veum, Hein de Wilde, (ECN/TNO), Vesna Bukarica, Lovorko Marić, and Marko Matosović (EIHP), Dario Di Santo and Enrico Biele (FIRE), Fabian Voswinkel (Fraunhofer ISI), Jean-Sébastien Broc, Mia Dragovic, Jan Rosenow and Niki-Artemis Spyridaki (IEECP), Romualdas Skema and Farida Dzenajaviciene (LEI), Lea Gynther and Ulla Suomi (Motiva)

With thanks to:

Gregor Thenius (AEA), Lea Gynther (Motiva Oy) and Dominique Osso (EDF R&D) for their valuable comments.

Project coordination and editing provided by Austrian Energy Agency.

Manuscript completed in August, 2018

This document is available on: www.epatee.eu

Document title	Evaluation into Practice: Lessons learnt from 23 evaluations of energy efficiency policies – Volume II: Background report
Work Package	WP3
Document Type	Public report
Date	29 August 2018
Document Status	Final version

Acknowledgments & Disclaimer

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

The EPATEE project

Several barriers limit energy efficiency policy evaluation. This results in a lack of quantitative data, and impedes evidence-based analysis required to distinguish effective from ineffective energy efficiency policies. EPATEE aims at tackling this problem by raising the capacity of policymakers and implementers. The project 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. Experience sharing is the core of the project. 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. For further information please visit our website: www.epatee.eu

Foreword

This volume is the background report of task 3.2 of the EPATEE project. It provides a complete view of the different points analysed in the EPATEE case studies about evaluations of energy efficiency policies and programmes. It gathers the main information collected in these case studies, enabling the readers to see the basis used to draw conclusions from the case studies. Some of these conclusions are presented in this report as “key messages”.

The analysis of this information is presented in the Volume I (Main findings) of the report about task 3.2 of EPATEE. The objective of separating the presentation of the work in two volumes was to keep the synthesis report (Volume I: Main findings) as concise as possible, while being as transparent as possible by providing the rough material used for the analysis in this report (Volume II: Background report).

Executive Summary

This report is the background report about the work done in task 3.2 of the EPATEE project (see *Foreword* for more explanations). The information presented in this report was collected from a set of 23 case studies of energy efficiency policies and programmes (listed in part 1). The following parts of the report present the different issues analysed in the case studies, the corresponding information collected and the messages drawn from them.

MESSAGE 01: Evaluation is not a burden, but an opportunity to strengthen policies and programmes.

This conclusion was drawn from the review of the reasons reported by stakeholders about why doing evaluation (section 2.1), showing the diversity of possible evaluation objectives, and above all providing a rich set of practical examples of evaluation use and lessons learnt from evaluations (section 2.2). This shows the added value that evaluation can bring to stakeholders.

MESSAGE 02: Evaluation priorities depend on its primary audience

MESSAGE 03: Evaluation helps to increase stakeholders' confidence, and thereby involvement in the policies (including financially).

These conclusions were drawn from the information presented in part 3 of this report. First a review of the evaluation management and requirements shows the diversity in the organisation and role of evaluation, and the importance not only of who commissions the evaluation, but also to whom the evaluation will be reported (section 3.1). Second feedback collected about the issues of confidence, legitimacy or credibility related to evaluation shows that stakeholders have confidence in the results if they trust the evaluators and their methods (section 3.2). Finally, examples provide insights about how to increase stakeholders' interest and confidence in evaluation (section 3.3).

MESSAGE 04: Monitoring and data collection are essential for making any evaluation possible.

MESSAGE 05: Selecting the most relevant data to collect is a continuous process.

Both conclusions are drawn from information presented in part 4 of this report. First a review of how monitoring and data collection are organised reminds the no-brainer that it is essential to plan data collection when designing or adapting the policy measures (section 4.1). Second feedback on how the data to be collected are selected shows that data collection procedures are fine-tuned over time (section 4.2). Then feedback provides a basis for experience sharing about the difficulties encountered (section 4.3) and examples of good practices about data collection and monitoring (section 4.4).

MESSAGE 06: Regular review and in-depth ex-post evaluations are complementary.

The analyses done in the case studies led to distinguish two main practices of ex-post evaluation: regular reviews (usually annual) and in-depth ex-post evaluations (usually on multi-year periods). Sections 0 and 5.2 details the methods used for annual reviews and multi-year evaluations respectively. Section 5.3 then gives an overview summarizing for which policy measures regular reviews and multi-year evaluations have been used, showing that most often both are used. Which supported the conclusion about their complementarity. Finally, section 5.4 provides practical examples of this complementarity.

MESSAGE 07: The choice of evaluation methods depends on evaluation objectives but also on practical constraints.

MESSAGE 08: There is no silver bullet. All methods include uncertainties. Comparing different methods helps assessing the robustness of the results, and getting a sound basis for decision-making.

Both conclusions are drawn from the information presented in part 6. Sections 6.1 and 6.2 gather the feedback collected in the case studies about the use of engineering methods and statistical methods respectively. This shows that there is no "perfect" evaluation method. All methods have their pros and cons, added value and limitations. Section 6.3 then provides examples where several methods have been compared, used to test the plausibility of the results or to get complementary analysis. In addition, section 6.4 looks at the sources of uncertainties identified in the different case studies.

MESSAGE 09: Analysing causality is a challenge, but essential to assess efficiency of policies.

The information presented in part 7 shows a diversity in the way to define or handle the concept of net or additional impacts, as well as in the methods used. Difficulties were also often reported, sometimes impeding the use of the method initially planned. This confirms that evaluating net impacts is a challenge. Stakeholders who tackled this challenge also reported that this is essential to assess and better understand the efficiency of the policies.

MESSAGE 10: good data is well-documented data.

In addition to be illustrated by a quote from one of the case studies, this conclusion comes from the difficulties encountered about clarifying the data collected when preparing the case studies (in terms of scope, unit, etc.). Sections 9.1 and 9.2 presents the different criteria used in the EPATEE case studies to document in a systematic way the data about energy savings and costs respectively. This shows the variety of metrics used to report energy savings, as well as the diversity in the scope of costs (when data about costs could be found). The data themselves are gathered in summary tables in Annex I.

MESSAGE 11: don't neglect discussion and communication about evaluation results.

This point was not directly in the scope of the EPATEE case studies, but it comes out from several of the interviews. So it was not systematically covered in all case studies. This will be further investigated in the next phase of the EPATEE project, in the task about how to integrate evaluation into the policy cycle. The first feedback collected in the case studies, and shown in part 10, provides a good starting point to include the communication issue in this topic of integrating evaluation into the policy cycle. It shows that the discussion and communication about evaluation results can be as important as doing evaluation.

The priority of the EPATEE project is to look at the evaluation of energy savings from energy efficiency policies and programmes. Therefore the analyses in the case studies were focused primarily on reviewing the data on energy savings and the methods used to evaluate them. Preliminary interviews and online surveys done by the project to assess stakeholders' needs and priorities highlighted their high interest in the evaluation of other indicators and impacts than energy savings.

Therefore the case studies were also used to review what types of other indicators or impacts were evaluated together with the energy savings: types of indicators to assess cost-effectiveness and cost-efficiency (section 8.1), impacts other than energy savings (section 8.2) and other aspects (process evaluation, customer journey, participants' satisfaction, market transformation, etc. ; section 8.3). It

should be noted that the EPATEE case studies did not intend to analyse the methods used to assess these other indicators or impacts, but to review which ones were evaluated.

Table of Contents

1 | List of EPATEE case studies

The information presented in this report was collected from a set of 23 case studies listed below. The case studies can be found on the EPATEE website: https://epatee.eu/case-studies

The examples presented along this report are not exhaustive. They do not reflect all the evaluation details or practices of the public bodies in charge of the policy measures. They are the ones that were highlighted in the interviews or that could be found in the evaluation reports. Direct quotes from the interviews are included with quotation marks.

When a case study is not included in the tables presenting information for the different points covered hereafter in this report, it is because no information about this point was found in the case study.

Table 1. List of EPATEE case studies.

Country	Sector	Type of instrument	Name of the measure	Done by
Austria	Industry and services	Financial	Environmental Support' (Umweltförderung im Inland)	AEA
Austria	Transversal	Policy mix	City Energy Efficiency Programmes of Vienna	AEA
Belgium (Wallonia)	Residential	Financial	Primes Energie (grants for energy renovation)	IEECP
Croatia	Services	Financial	Energy renovation of public sector buildings programme	EIHP
Croatia	Residential	Information/Education /Training	Individual heat metering in multifamily buildings	EIHP
Denmark	Transversal	Market-based	EEO scheme	IEECP
Finland	Services	Cooperative	Energy Efficiency Agreement for Industries	Motiva
Finland	Services	Information/Education /Training	Voluntary energy audits for municipalities	Motiva
France	Transport	Cooperative	Voluntary agreement for freight companies	ADEME
France	Transversal	Financial	"Future Investments" programme	ADEME
Germany	Industry and services	Cooperative	Energy Efficiency Networks Initiative (Initiative Energieeffizienznetzwerke)	Fraunhofer ISI
Germany	Transversal	Financial	Energy Efficiency Fund	Fraunhofer ISI
Ireland	Residential	Financial	Better Energy Homes	IEECP
Italy	Transversal	Market-based	White Certificates Scheme	FIRE
Lithuania	Residential	Financial	Renovation programmes with EU funding	LEI
Netherlands	Industry, agriculture and services	Cooperative	Multi-year agreements in the industry	ECN

Country	Sector	Type of instrument	Name of the measure	Done by
Netherlands	Residential	Financial	Subsidy scheme for housing corporations in Amsterdam	ECN
Netherlands	Transport	Fiscal	Purchase tax on passenger cars	ECN
Nordic Countries	Residential	Legislative/normative	Nordsyn (market surveillance for the EcoDesign Directive)	IEECP
UK	Residential	Market-based	Supplier Obligations	IEECP
UK	Residential	Financial	Warm Front	IEECP
US	Energy sector	Market-based	Auctions for capacity markets in New England	IEECP
US	Residential	Financial	Weatherization Assistance Program	IEECP

2 | Evaluation: why and what for?

The main conclusion from this part is:

MESSAGE 01: Evaluation is not a burden, but an opportunity to strengthen policies and programmes.

This is illustrated in the information collected below. First about the reasons why doing evaluations, showing the diversity of possible evaluation objectives (section 2.1). Second about practical examples of evaluation use and lessons learnt from evaluations, showing the added value that evaluation can bring to stakeholders (section 2.2).

2.1 Feedback about why doing evaluation

2.1.1 Accounting for and reporting about results, effectiveness and efficiency (summative dimension of evaluation)

Table 2. Objectives of the evaluation reviewed (about accountability/summative dimension).

Case study	Evaluation objectives / quotes
[AT] Environmental Support scheme	"the programme is evaluated every three years by external evaluators. The requirement for evaluation is set by law." "It is used to report to the federal parliament on the effects of the scheme" Focus on ecological, budgetary and economical effects
[AT] City Energy Efficiency Programmes of Vienna	"There is an obligation to report on the progress of the programme every 3 years to the city council. (…) The evaluation gives feedback about the progress of the implementation of the single measures and it helps to get quantitative information on energy savings." More in details, the summative objectives of the final evaluation done in 2015 were: to assess and document the implementation progress and overall; to quantify the overall impact of the measures to reduce energy consumption growth; to comply with the reporting requirements of the Energy Services Directive (2006/32/EC).
[BE] Primes Energie	Reporting energy savings to the European Commission Providing evidences about results as inputs for decision-making about the policy measure (particularly its budget)
[CR] Energy renovation of public sector buildings	Detailed evaluation, i.e. an analysis focusing on cost effectiveness, is performed as new plans are being developed. Therefore, at least a rudimentary cost-effectiveness evaluation is integral part of every new plan for the new period of the programme (e.g. 2016 – 2020).
[CR] Individual heat metering in multifamily buildings	Assessing actual energy savings and cost-efficiency of the actions (installing individual heat cost allocators)
[DK] EEO scheme	"these evaluations aim to investigate the satisfaction of the stakeholders (energy distributors, contractors, end-users), the impacts in terms of energy savings and the cost-effectiveness of the scheme (from a society point of view)." This is "important to know if the scheme meets its objectives (…) and to know if the energy distributors are using cost-effective approaches to deliver the energy savings, which is an essential criteria for the cost recovery mechanism". Energy distributors can indeed recover on the energy distribution tariffs the costs incurred to achieve their targets, under approval by the regulatory agency (DERA). Assessing additionality is the key to ensure that the scheme delivers a net benefit to all end users.

Case study	Evaluation objectives / quotes
[FI] Energy Efficiency Agreement for Industries	"We need to be constantly aware of the progress made. Good results ensure that actions are sustained." A well-functioning monitoring system for the agreement scheme has had a central role in revealing the results, create trust and credibility among all parties (not just the government) and in achieving long-term top-level commitment.
[FI] Voluntary energy audits for municipalities	"The overall performance of the scheme is continuously evaluated."
[FR] Voluntary agreement for freight companies	"The reasons of this evaluation were the needs to get an overview of the impact of this innovative and dynamic scheme which was in operation for many years (the scheme was launched in 2008), and to prepare its future." An evaluation diagnosis was made to assess what data were available and what data could be collected. Then it was decided to focus the evaluation on the following aspects for its summative part (impact evaluation): • the effectiveness (objectives reaching) • the efficiency (ratio cost-benefit) • impact measurement (unexpected effects)"
[FR] "Future Investments" programme	Complying with the EU rules for State aids supporting environmental protection and energy objectives, and the EU obligation to evaluate large state aid programs (annual expenses ≥150 M€) Contractual duty for ADEME and the State to determine the feasibility of an econometric evaluation of the programme The objective of the econometric evaluation was to report on the direct and indirect impacts of the aid scheme
[DE] Energy Efficiency Networks Initiative	Monitoring the achievement of the targets set by each network, and estimating the overall impacts of the scheme (in terms of primary energy and CO_2e savings)
[DE] Energy Efficiency Fund	"Evaluations serve as a means of justification for actions taken. The Bundesrechnungshof (German Federal Court of Auditors) often uses evaluation reports for their judgement about the usage of public funding." The goals of the evaluations are laid out in the methodology report setting the common methodological principles to be followed in each evaluation.
[IE] Better Energy Homes	"BEH is the biggest grant scheme for energy efficiency. The government therefore pays a particular attention to it. And the Public Spending Code requires that every scheme beyond a given threshold of annual public expenses shall be reviewed periodically. The first part of the evaluation (Cost-Benefit Analysis, CBA) was thus done to have an evidence base for the discussions with the Ministry of Finance, in particular about value for money and contribution to national targets. This was needed as SEAI is accountable to the government, and more generally to society." "Evaluation is crucial when public money is spent. This is a matter of responsibility and transparency. We need to know how much money is spent and how. And what impacts are achieved."
[IT] White Certificates Scheme	Evaluation is primarily done at the project level to assess the amount of additional energy savings that can accounted for issuing white certificates. Then an annual review of the scheme is done to see if the results are in line with the targets. "The role of evaluation is crucial to assess the eligibility of an energy efficiency project and the amount of WhC required. In particular the evaluation process is even more determinant to evaluate the additional energy savings (from baseline) and the affordability of the measurement campaign proposed by investors to demonstrate the energy saving related to the energy efficiency projects."

Case study	Evaluation objectives / quotes
[LT] Renovation programmes with EU funding	"Typically the following evaluation is implemented for the energy efficiency policies in Lithuania: • Ex-ante evaluation is performed mainly as some technical evaluation for specific measures • Ex-post evaluation is performed partly and not for real lifetime of energy efficiency measure" Evaluation is mostly based on the monitoring of the results achieved by each policy. Complementary studies have been done to assess ex-post the actual energy savings achieved by samples of projects.
[NL] Multi-year agreements in the industry	"The ex-post evaluation of the LTA3 aimed at investigating if the scheme meets its objectives and to identify how the scheme could be improved." The ex-post evaluation also aimed to assess what were the implementation costs for government and industry in relation to the benefits.
[NL] Subsidy scheme for housing corporations	The objective of the evaluation done by the Amsterdam audit office was to assess the actual impacts of the scheme (in terms of energy and CO_2 savings).
[NL] Fiscal incentives for cars	"The value and role of the evaluation of the bpm fiscal measure was to get a better picture of the effectiveness of the fiscal measure." Other aspects that were evaluated include: • use of the purchase tax measure to stimulate the sale of clean and efficient cars; • leakage effects of the Dutch tax measure; • test procedures for analysing CO_2 emission (difference in test lab results vs. results in the field).
[Nordic Countries] Nordsyn	"The main role of market surveillance is to make sure that savings are occurring and that all market actors are aware of the requirements." Then the evaluation was done to assess if the market surveillance activities were cost-effective by comparing the energy savings that would be lost without market surveillance and the costs of the market surveillance activities.
[UK] Supplier Obligations	"There is also continuous tracking of impacts of scheme in terms of actions and costs" Key objectives of the monitoring, verification and evaluation activities are to monitor target achievement, update the values used for deemed scores (energy savings ratios per standardised types of action), assess the value for money of the scheme.
[UK] Warm Front	"The general context was the increasing issue of fuel poverty in the early 2000's, and particularly excess winter deaths that were estimated to be related to fuel poverty (about 20 000 per year at that time). The driver for this large research project was the government's interest in the health impact of energy efficiency and the potential for the Health Departments involvement in the scheme including contributing to its costs. The aim was to evaluate the health benefits, which could save lives. This assumption was therefore to be tested as if the Warm Front interventions were a kind of medicine. This meant that the general evaluation question was to investigate if the scheme was cost-effective in delivering the expected health benefits. (…) Evaluating the Warm Front scheme in this view was a particular case. And the research team managed to convince the administration that it was also needed to look at intermediate aspects and indicators, such as changes in indoor temperatures and fuel consumption. That is how one of the studies included in this research project could look at the energy-related aspects."
[US] Auctions for capacity markets	"We need to ensure that resources can provide capacity when needed to meet projected resource adequacy needs."
[US] Weatherization Assistance Program	"The primary evaluation objective is thus to assess the energy savings then converted into bill savings, and compare them with program costs" Cost-effectiveness is evaluated against the SIR (Savings to Investment Ratio), and also compared to previous periods that provide a benchmark for next years.

2.1.2 Examining what works, what does not work, looking for improvements and questioning new ideas (formative dimension of evaluation)

Table 3. Objectives of the evaluation reviewed (about understanding/formative dimension).

Case study	Evaluation objectives / quotes
[AT] Environmental Support scheme	"It is used internally by the competent Ministry and the management body to get hints not only on the results/effects but also on possible future requirements and focus topics of the scheme. The evaluation of the scheme tries to go a little more into detail of why certain things happened in order to understand the underlying dynamics and be able to take adequate measures to solve these issues." Evaluation of the processing time (improving the effectiveness of the processing of the applications)
[AT] City Energy Efficiency Programmes of Vienna	"the evaluation is important for the scheme from an operational point of view" More in details, the formative objectives of the final evaluation done in 2015 were: to include a description of each measure included in the SEP and an analysis of the development of relevant legal framework conditions (at European and Federal levels); to evaluate the existing catalogue of measures based on objective and comprehensible criteria; to provide recommendations for the design of a successor program ("SEP 2030").
[BE] Primes Energie	improving consistency and management/administration of the policy measures
[CR] Energy renovation of public sector buildings	understanding financial challenges and organizational issues
[CR] Individual heat metering in multifamily buildings	feasibility and cost-efficiency of installing individual heat cost allocators in Croatia, analysing under which conditions this would be feasible
[DK] EEO scheme	investigate whether the rules of the scheme were appropriate, the level of satisfaction of the stakeholders (obligated parties, end-users, etc.) This is "is important to know how the scheme can be improved in practice." "The ex-post evaluations are used to complement the monitoring of the scheme when preparing a revision of the agreement for the scheme." "Policies are living creatures and need to be adjusted periodically to take into account changes in context, markets, policy priorities, etc. A timely evaluation can provide the necessary basis for this"
[FI] Energy Efficiency Agreement for Industries	reason for carrying out the evaluation in 2004-2005 was that results and ideas for further improvement could be used in the planning of future activities
[FI] Voluntary energy audits for municipalities	"If there are problems, we need to know where those are. It is another question if we can interfere, but we must know and understand the situation."
[FR] Voluntary agreement for freight companies	In addition to the objective of assessing the impact of the scheme, the evaluation was also meant "to prepare its future." An evaluation diagnosis was made to assess what data were available and what data could be collected. Then it was decided to focus the evaluation on the following aspects for its formative part (process evaluation): • the intrinsic relevancy of the scheme (coherence between objectives & the sector needs and evolution) • the internal coherence (adequacy between the scheme and means) • the external coherence (adequacy between the scheme and other programmes)

Case study	Evaluation objectives / quotes
[FR] "Future Investments" programme	the survey provided qualitative information on project management and the effects of aid on innovation, partnerships, collective learning, and commercial and technological opportunities. This method was designed to answer the following evaluation questions: 1) To what extent are the proposed support measures best adapted to needs in the area of innovation (relevance analysis)? 2) To what extent the programme has produced the anticipated effects (effectiveness analysis)? 3) Is the programme cost-effective?
[DE] Energy Efficiency Networks Initiative	Getting a feedback and assessing the satisfaction of the scheme administrators and participants
[DE] Energy Efficiency Fund	Satisfaction of the beneficiaries and administrators, with a particular focus on satisfaction with the provided information, processing times and support offered by the scheme Getting recommendations to improve the schemes and new ideas for policy making
[IE] Better Energy Homes	"Qualitative analysis is also essential, for example to know how the participants feel about the improvements of their dwelling. This should be combined with the quantitative impact analysis, in order to understand how to promote the scheme."
[IT] White Certificates Scheme	The regular monitoring of the scheme is also aimed at fine-tuning the rules of the scheme. The scheme has also been subject to various studies, especially about the trading mechanisms and the market behaviours according to the changes in the rules of the scheme.
[LT] Renovation programmes with EU funding	National Audit Report done in 2010 to assess how the programme worked and could be improved. The Housing Energy Efficiency Agency commissioned a survey to investigate satisfaction of flat owners and residents after renovation works. A research (master thesis) was done in 2012 to identify the main barriers to the renovation of apartment blocks and suggest recommendations to overcome them (mostly based on benchmarking of experiences in other countries).
[NL] Multi-year agreements in the industry	In addition to assess if the objectives were met, the ex-post evaluation of the LTA3 also aimed at identifying how the scheme could be improved. "Earlier ex-post evaluations, i.e. for LTA1 and LTA2, were done for two main reasons, namely: (i) development of new insights and (ii) the connection to European directives."
[NL] Fiscal incentives for cars	The objectives of the evaluation by PBL was to get a better insight in the effectiveness, opportunities and barriers of different policy instruments.
[Nordic Countries] Nordsyn	The evaluation (Effect study) also aimed at making suggestions to improve the strategy of appliance tests.
[UK] Supplier Obligations	"The purpose of evaluation is to evaluate the scheme, consider learnings and implement changes in next phase." Monitoring and evaluation aims at fine-tuning the implementation of the scheme and identify needs for updates in the scheme settings.
[UK] Warm Front	The objective of the final process evaluation was that lessons could be learned to inform the delivery of future energy efficiency schemes targeting the fuel poor (in particular for the Energy Company Obligation). This evaluation investigated, through a combination of quantitative and qualitative research methods, the strengths and weaknesses of scheme processes.
[US] Auctions for capacity markets	For utility EE portfolios that bid into the ISO-NE forward capacity market, they are typically also subject to other types of evaluation than the peak savings verification. While these evaluations often do not affect the ex post peak demand reduction verification, they are valuable for understanding how the design and implementation of EE schemes can be improved and forecasting market trends and system operational needs.

Case study	Evaluation objectives / quotes
[US] Weatherization Assistance Program	"DOE-WAP (WAP team of the Department of Energy) sees value in evaluation. Not only how it is important to justify the program, but also to improve it and guide policy decisions." The part of process evaluation explored how the weatherization network delivers services, evaluated how service delivery under WAP compares to national standards, documented how weatherization staff and clients perceive service delivery.

2.2 Practical examples of the added value of evaluation: evaluation use and lessons learnt

Table 4. Examples of evaluation use and lessons learnt reported by the stakeholders interviewed for the EPATEE case studies.

Case study	Quotes / comments
[AT] Environmental Support scheme	"Apart from showing the effects of the scheme the evaluation helped to adapt the scheme and its rules to make it more practical and effective." Willingness for development and adaption of the policy according to the results of the evaluation was also pointed as a crucial point to enable a good evaluation and then evaluation use. Evaluation is used to detect new trends and changes by making comparisons between periods. Practical example of improvements made after evaluations: - requirement of a quality management system for district heating projects, avoiding oversizing and thereby decreasing investment costs needed; - introduction of maximum ceilings of aid per tCO_2 - introduction of lump sum for small projects - digital application tool to facilitate data collection and processing of applications - improvement of the database gathering the data collected through the application process and monitoring of the scheme
[AT] City Energy Efficiency Programmes of Vienna	"The evaluation was very important and helpful in order to optimise the programme management and to focus on the important tasks and measures. The interim evaluation reports included recommendations for the next implementation period. So it was possible to focus on specific important measures and to learn from experiences made." One of the objectives of the final evaluation done in 2015 was to provide recommendations for the preparation of the next programme, SEP 2030. The recommendations were about strategic aspects (e.g., number and targeting of the measures), as well as on operational aspects, including suggestions to improve monitoring & evaluation.
[BE] Primes Energie	"An example of decision is that sun protection/shading was removed from the list of eligible actions from 2015, based on the low effectiveness assessed and the fact that priorities were set on building envelope and heating system's improvements." "another lesson from this recent experience is that it is very difficult to predict quantitatively the effects of a change in the characteristics of a measure (for ex., incentive rates, eligibility conditions, application process). It is obvious that the incentive rate has an impact on the attractiveness (and thereby effects) of the scheme. But it is much less obvious that an elasticity could be assessed between the incentive rate and the number of applications, and that this elasticity would be linear. So any change in a measure should be done very carefully, analysing the pros and cons." "About additionality, another lesson learnt is that trying to limit free-rider effects may lead to unexpected negative effects (for ex., if the decrease in the applications is stronger for households who would be the most in need)." Better understanding of what makes an incentive attractive (not only incentive rate, but also whether the application process is easy and in line with the timing of renovation projects) The chronological analysis of the data about grant applications for energy renovation works showed that households were very reactive to the changes in the requirements and rates of the incentive.

Case study	Quotes / comments
	The evaluators therefore pointed the risk of too sudden changes that affect the demand for works, and thereby business cycles. Whereas a stable market environment is needed for companies to be able to invest in new technologies and training. Frequent changes in grant conditions also make it difficult for companies to inform correctly households. In parallel, a detailed analysis on data about roof insulation pointed a trend towards materials with higher performance. The econometric analyses showed that participants with higher income had higher probability to do works with natural or super-efficient materials compared to households with lower income. The evaluators then recommended to use different grant rates depending on income classes.
[CR] Energy renovation of public sector buildings	Results of all those analyses are then fed into the plan for a new round of energy efficiency policy measure(s). Future policy measures are then analysed in four scenarios under different organizational and financial setting. One of the recommendations was the development of (a) guarantee instruments, (b) equity instruments, or (c) new credit line.
[CR] Individual heat metering in multifamily buildings	The cost-effectiveness, and therefore attractiveness for the consumers, was shown to be highly dependent on the price of heat (which varies among cities) and the type of buildings. Recommendations to improve technical implementation of the actions and billing calculations (formula for heating cost allocation), to increase customer protection, and to revise the subsidy and financing scheme.
[DK] EEO scheme	"All the recommendations of the ex-post evaluations have not necessarily been implemented. But they have been discussed with the obligated parties, and many of them have been used, either directly and with some further adaptations" "For example, simple prioritisation factors have been adopted after the first evaluation (2008), in order to take into account the differences in lifetime per action type, and to favour actions with longer lifetimes. Another example is the introduction after the second evaluation (2012) of a rule that actions with a payback time of less than 1 year could not receive a grant, as this was raised as an additionality issue." "In the 2012 evaluation of the energy efficiency obligation scheme, we found that information campaigns carried out by the obligated parties were difficult to distinguish from ordinary company PR campaigns. Information campaigns without strong documentation of net impact were subsequently removed from the eligible list of energy efficiency actions." + CFL and appliances removed from the scheme from 2010
[FI] Energy Efficiency Agreement for Industries	"The success factors of this well-working policy measure have been good monitoring and evaluation, strong results and communication of results." A well-functioning monitoring system for the agreement scheme has had a central role in revealing the results, create trust and credibility among all parties (not just the government) and in achieving long-term top-level commitment. Following the discovery that small organisations and SMEs in particular need a support person who would provide them with both general information on the energy efficiency agreement scheme and specialist advice on their own sector, energy advice to SMEs was implemented in the 2008-2016 period. Companies were supported in their communications by providing more communication services from coordination and the web-pages were developed to include, e.g., case studies on good practices. A large new effort was the shift in reporting from Excel to an on-line system.
[FI] Voluntary energy audits for municipalities	"The role of continuous evaluation was critical during the first years and it has been very important afterwards." "Concerning energy audits and voluntary agreements it is so normal thing to have that we don´t question it. It is a sort of guarantee that what we do has a clear and unquestioned justification." "the provision to nominate two auditors (one for heating, ventilation and air conditioning (HVAC) and one for electrical systems) came from the fact that using only one auditor delivered nearly zero savings from the "other system". Also, constant low level of savings compared to the average is an indication that there may be a quality problem" "Information collected is used also in quality control, in preparing lists of energy saving actions, marketing of energy audits and in reporting of savings to the EU." "When we started promoting ESCOs we needed to ensure that we do not create competing policy

Case study	Quotes / comments
	measures. We have also seen an interaction between the subsidies for investments and audit volumes. If subsidies are taken away, that will lower the interest in audits." "We used our knowledge gathered thru our long history of M&V (1994->) of voluntary energy audits when we planned M&V for the mandatory audits"
[FR] Voluntary agreement for freight companies	"The evaluation led in 2013 has highlighted in particular: • additional benefits of the scheme on pollutants emission reduction (that led to the enlargement of the scope to pollutants such as NOx and particles matters); • the need to reinforce the reliability of data transmitted by transport operators by introducing a labelling system based on data checking and the control of the objectives achievement (that led to the launch in 2015 of a label including an audit procedure); • the need for reporting procedures improvement in particular between central and regional administration." "All evaluations supported by ADEME are reviewed in order to monitor the implementation of recommendations but operational decisions to ensure this implementation are not necessarily taken into account." The part of process evaluation provided complementary insights about the effectiveness of the scheme and how it could be improved. For example, this enabled to identify reasons why companies were or were not joining the programme. It also highlighted a lack of information transmission between the regional department of ADEME and the headquarter partly because of the lack of reporting tool. The evaluation also recommended to implement a voluntary label to improve the reliability of the monitoring and verification, and the effectiveness of the scheme (by requiring labelled participants to do external audits of their procedures). This has been implemented.
[FR] "Future Investments" programme	The qualitative mid-term evaluation of the programme has helped readjustment of evaluation processes & some improvements in internal procedures for investigating and contracting projects. The mid-term econometric evaluation performed over the period [2009-2013] has enlightened in particular many difficulties in impacts observation mainly due to the too short duration of the observation, sampling losses, non-reliable data transmitted by projects leaders (due to memory losses and lack of time & incentives for them to search for accurate data). These issues led ADEME to the design tools facilitating data collection for the programme monitoring and evaluation especially for beneficiaries of refundable aids. This mid-term evaluation has also led to more qualitative objectives enabling to explain the projects status: immediate effects (and not only long-term effects) and to understand precisely the reasons of innovations success and failures, and especially from business point of view. Last but not least the mid-term evaluation has also confirmed the difficulty in operationally implementing econometric methods and the difficulty in obtaining robust figures. However, it has helped understanding the causal processes leading (or not) to technological and/or commercial successes
[DE] Energy Efficiency Networks Initiative	"Networks are considered a crucial instrument for climate protection that requires evaluation and consequently requires participants to hand in proofs of savings" "While the target number of 500 networks may not be reached until 2020, the 75PJ (20,8 TWh) savings target is within reach making the networks initiative among the most successful policies in the NAPE (National Action Plan for Energy Efficiency) policy set" "One criticism is the low number of requirements and the varying network runtimes. This way networks adjust very much to the necessities of the participants, which can on the one hand reduce barriers to participate and lead to more savings, but on the other hand, it gives away control about the governing of the network initiative." "The feedback from both sides [administrators and participants] is almost exclusively positive once they are part of the programme. That is good news for the continuation of the initiative."
[DE] Energy Efficiency Fund	"Evaluations serve as a means of justification for actions taken. The Bundesrechnungshof (German Federal Court of Auditors) often uses evaluation reports for their judgement about the usage of public funding." "Recommendations are frequently used for the implementation and modification of the energy efficiency support strategy. They are generally given high importance in decision-making about energy efficiency policies and serve as intellectual input for policy ideas. (…) However, recommendations that are not possible to implement or are politically not opportune, may be left

Case study	Quotes / comments
	out of consideration." "About the evaluation of current industry related energy efficiency programmes. The evaluation strongly supports the strategy of technology-independent support schemes. They can bundle different technology categories in one programme, reduce precise technological requirements and therefore reduce administrative costs and barriers for potential beneficiaries. That is why the programmes should serve as a benchmark for other policy measures." "In the exhaust heat project, it gave the important insight, that the concepts from energy auditors that serve as the basis for the granting of aid should be more comparable. More detailed requirements should be formulated."
[IE] Better Energy Homes	"The ex-post evaluation was very well accepted by the Ministry. Indeed the evidence brought by the evaluation changed their perception of the scheme. There were no more questions about the rationale or interest to implement this scheme. At the opposite, the questions were about how to make the scheme grow." "One of the key results of the evaluation was that the energy savings impacts differed depending on the initial level of energy performance of the dwellings (based on the Energy Performance Certificates' classes). We observed higher comfort taking in dwellings that were the least energy efficient before works. This argued in favour of promoting more comprehensive renovations." "Another study is currently done to review each technology eligible to the scheme and the relevance to support it with public funds. The objective is to update the eligibility criteria and grant rates" "The comparison between the engineering estimates and the billing analysis was used to update the standard values of energy savings per action and dwelling types used to monitor energy savings from the data registered from the grant applications"
[IT] White Certificates Scheme	"Monitoring WhC evaluation process is really important for detecting the effective achievement, reasonableness and rationality of the goals settled during the design of the energy efficiency policies." The evaluation of the scheme is important to know its contribution to the Italian energy efficiency strategy: "The amount of energy savings generated by white certificates are higher than all the other energy efficiency measures active in Italy (fiscal rebate, heating account, energy label etc.). Currently the scheme is characterized by the lowest cost/saving ratio."
[LT] Renovation programmes with EU funding	The National Audit Report of the programme done in 2010 identified weaknesses and made recommendations that led to improvements made after the 2007-2013 period: • Municipalities were instructed to draw list of the worst-performing buildings (to ensure a better targeting); • Municipalities have appointed renovation administrators (to facilitate the process of the renovation projects); • Administrators borrow on behalf and for the need of apartment owners (to facilitate the financing of the renovation projects); • Loan remains off balance sheet for Administrator; and • Amendments to the legal basis related to heating bill compensations were adopted on 1 June 2013: where a community decides to renovate the multi-apartment block, those low income families who have declined participation in the decision-making process would receive 50% smaller compensation for the heating bills during the proximate heating season and no compensation from the next heating season until the block renovation project is completed but no longer than 3 years.
[NL] Multi-year agreements in the industry	Ex-post evaluations are used by NL Agency to improve the schemes to promote energy efficiency. For example, the conditions to comply with the LTA (and thus get the tax exemptions) were tightened in 2015. "Benchmarks allow evaluators to better judge what is efficient."
[NL] Fiscal incentives for cars	"The measure is quite effective from the standpoint of the Dutch government. However, car companies may choose to sell more fuel efficient cars in the Netherlands (because of the favourable market conditions with the reduced purchase tax) and sell less of these fuel efficient cars in other European countries, initiating a "waterbed" effect. This is possible since manufacturers have to meet an EU wide efficiency target (rather than national targets). Controlling CO_2 reduction with the fiscal measure of CO_2 differentiated purchase tax (bpm) is

Case study	Quotes / comments
	effective, as long as boundary conditions (such as the waterbed effect) are known and can be managed. Another lesson is the fact that the bpm measure was an open ended measure for a 4 year period. This means there is no limit to the financial scope of the measure (as opposed to e.g. a subsidy measure)."
[Nordic Countries] Nordsyn	"One of main lessons is that close cooperation among MSAs [market surveillance authorities] and policy experts of different countries creates knowledge and possibility for all to advance. Nordic countries with our Nordsyn programme is a quite close group of professionals that do not hesitate to ask questions and discuss difficult issues with one another. With over 50 regulations in place for Ecodesign and Ecolabelling, such cooperation and synergy is a key in understanding all the policy, and implementing its requirements at the least cost. The countries MSAs and policy professionals work together on regular basis, motivating each other to do more, and enabling us, for example, to encompass a bigger share of the market than any of us could do on our own. " Results of the evaluation confirmed the cost-effectiveness of Nordsyn actions (market surveillance cooperation), demonstrating with a cost-benefit analysis the impact of implementing market surveillance. The Nordic Council of Ministers thus adopted a budget to continue Nordsyn over 2016-2017. The evaluation also made suggestions on how to better choose products to test. One of ways is to choose group of products more sensitive to more strict surveillance (e.g. larger share of savings when the non-compliance factor is increased). For this reason, a sensitivity analysis was performed for each group of products, to see how the savings react when non-compliance is increased by 1%. They found that the electric motors, standby and lighting would be priority groups of products to focus market surveillance efforts.
[UK] Supplier Obligations	"We identified a number of lessons learnt: a) We consider the work with trusted bodies such as local authorities as very important. We have introduced a new flexibility options for obligated parties to deliver a share of energy efficiency actions through local authorities. b) There is a need for transition agreements between periods to phase-in any changes and ensure continuity. c) Data sharing is important, particularly for targeting fuel poor households - BEIS is currently working towards sharing data from the Department for Work and Pensions with energy suppliers. d) The quality of the installed actions is very important and not always considered. We commissioned the so-called Bonfield review looking into how to ensure highest quality standards. We also established the National Energy Efficiency Data-Framework (NEED) to check what actions are actually delivering. Ofgem also carries out technical monitoring and spot checks to ensure quality standard are met. e) We learned that there is a need for a long-term framework including carry-over options (if savings targets are exceeded) and longer timeframes providing certainty for industry. f) We recognised the need for a better understanding of the costs to obligated parties and implemented a data requirements for suppliers to report cost. g) Data on consumer and third-party contributions to the cost is relatively scarce and further data will need to be collected. h) There is a need for a better understanding of fuel poverty impacts of the Supplier Obligation and impact research is under way."
[UK] Warm Front	"It is interesting to note that it is now possible to get doctors in effect prescribe a new boiler. This happened because health benefits from this type of intervention could finally be demonstrated to be cost-effective by NICE (National Institute for Health and Care Excellence) by evaluating data collected from a range of trials like Warm Front." "Overall, Warm Front measures – insulation and better heating systems - had a positive impact on improving mental health, improved thermal comfort and internal temperatures which in turn are linked to excess winter deaths, and lowered relative humidity. Through fuel switching, fuel costs were reduced but energy use did not reduce. About the energy-related aspects, the installation of central heating, although more efficient, increased temperatures which increased energy use. Also, draught stripping which should have reduced the heat loss was offset by increased air infiltration during the boiler installation due to new holes being drilled.

Case study	Quotes / comments
	This led to the recommendation that boiler replacement and other interventions should be done before interventions on airtightness. More generally, the evaluation confirmed that health impacts were larger than impacts on fuel consumption. The main improvements for the participants could indeed be found in terms of higher indoor temperature and better comfort."
[US] Weatherization Assistance Program	"There were two big results from the 1989 evaluation: • Savings in Southern states were found to be lower than in Northern states: it was thus decided to put efforts in better understanding this, which led to promote more baseload measures (e.g., lighting, refrigerators) as well as measures on air conditioning. • The evaluation recommended that states and Sub-Grantees should use computerized audits instead of predefined lists of actions. There was thus a switch from a generic to a more specific approach. This made that auditors spend more time in the house and thus better see what is really needed. The results of the latest evaluations have been published in 2014-2015. So their use is still under progress. One result was that radon level was increasing after weatherization. A follow-up study is being done to better understand this phenomenon. Provisions have already been taken about ventilation and quality of works, and especially covering up the ground and foundations (to prevent radon coming from the ground). It should be noted that WAP was changing while evaluations were done in 2008-2011. For example quality aspects became more important in both, the program and its evaluation. In the ARRA period, Standard Word Specifications (SWS) were introduced as an effort to document the proper way to install each type of measure. Then certifications were required for auditors, crew leaders and quality controllers. In parallel, the evaluation investigated the quality of the works and quality assurance processes. There will soon be a new evaluation to see the impacts of the new provisions about work quality. The results of the previous evaluations will provide the baseline for it. Another result of the last evaluations was that lower energy savings were found in mobile homes. Then DOE's Building Technologies Office is doing further investigations about retrofitting mobile homes to see how to improve this, for example for attic insulation. Another point is NEBs (Non-Energy Benefits). They have always been an important topic for WAP. The methodology tested during the last national evaluations is now used by the states for their own evaluation, which should bring more data and evidence about NEBs."

3 | Evaluation: who is involved and why?

The main conclusions from this part are:

MESSAGE 02: Evaluation priorities depends on who the primary audience is

MESSAGE 03: Evaluation helps increase stakeholders' confidence, and thereby involvement in the policies (including financially).

This is illustrated in the information collected below. First about the evaluation management and requirements, showing the diversity in the organisation and role of evaluation, and the importance not only of who commissions the evaluation, but also to whom the evaluation will be reported, and at what stage of policy implementation/maturity the evaluation is done (section 3.1). Second with feedback collected about the issues of confidence, legitimacy or credibility related to evaluation, showing that stakeholders have confidence in the results if they trust the evaluators and their methods. And sometimes raising the issue of what "independent" can mean when speaking of "independent" evaluation (section 3.2). Finally, examples provide insights about how to increase stakeholders' interest and confidence in evaluation (section 3.3).

3.1 Diversity of evaluation management and purposes

Table 5. Who is involved in the evaluations and related evaluation purposes or expectations.

Case	Stakeholders and roles	Evaluation purposes/expectations
[AT] Environmental Support scheme	Supervision by the Federal Ministry for Sustainability and Tourism that performs annual reviews and commissions an external evaluation every three years Management + M&V by Kommunalkredit Public Consulting Ex-post evaluation by a team of contractors (external) commissioned by the ministry	Reporting to the Federal parliament and the Federal Ministry of Finance who are responsible of the formulation and overall budget of the scheme Therefore, for the Ministry for Sustainability and Tourism, the first evaluation objective is to show the impacts that the programme has for the participants, particularly when reporting to the Parliament. Hence a focus on ecological, budgetary and economic effects. The external evaluation also investigates how the administration of the scheme could be improved. Evaluations are first done for the needs of the Ministry (own purpose). Other energy savings calculations are done in parallel to report results for EED article 7, which creates an additional burden.
[AT] City Energy Efficiency Programmes of Vienna	Programmes implemented by various City departments Coordination body at the City of Vienna, in charge of centralising the data and supervising the reviews and evaluations. Monitoring done by each City department for the schemes they are in charge of. Ex-post evaluation commissioned every three years by the Coordination body to external evaluators.	Obligation to report on the progress of the programme every 3 years to the city council as well as to the national monitoring body. Being able to have a view of the implementation and achievements of each of the about 100 measures of the programme, and to assess the overall progress made. Identifying opportunities of improvements.

Case	Stakeholders and roles	Evaluation purposes/expectations
[BE] Primes Energie	Management and monitoring-evaluation of the scheme done by the Energy Department, but by two distinct services. Review done by the Court of auditors (of a policy portfolio that includes Primes Energie). External evaluation (by public institute for evaluation and statistics) commissioned by the government (of the umbrella policy framework that includes Primes Energie).	At policy measure level, objective = monitoring and reporting of energy savings At policy portfolio level, objective = improving consistency and management/administration of the schemes At umbrella framework level, objective = assessing the effectiveness of the different measures, notably in terms of socio-economic impacts on the private demand and offer for sustainable buildings (focusing on renovations), and more specifically about employment effects
[CR] Energy renovation of public sector buildings	Monitoring with SMiV platform (managed by the Energy Efficiency Authority). Internal reports made by the Environmental Protection and Energy Efficiency Fund (FZOEU). Evaluation supervised by the Energy Efficiency Department of the Ministry of Environmental Protection and Energy.	Analysis focusing on cost effectiveness, when new plans are being developed. Understanding financial challenges and organizational issues.
[CR] Individual heat metering in multifamily buildings	Monitoring with SMiV platform (managed by the Energy Efficiency Authority) First evaluation commissioned by the Ministry of Economy. Second evaluation commissioned by the Ministry of Environmental Protection and Energy.	Evaluation objectives focused on the feasibility and cost-efficiency of installing individual heat cost allocators in Croatia, analysing under which conditions this would be feasible.
[DK] EEO scheme	Monitoring by Danish Energy Agency (DEA), based on data reported by the obligated parties, with further verifications and controls. External evaluation commissioned by DEA at the end of each period Obligated parties consulted when defining the evaluation questions, and intermediate discussions with the evaluators during the evaluation.	Evaluation objectives: investigate the satisfaction of the stakeholders (energy distributors, contractors, end-users), the impacts in terms of energy savings and the cost-effectiveness of the scheme (from a society point of view), providing a factual basis to prepare the next period. Assessing cost-effectiveness is particularly important due to the cost recovery mechanism (even if the impact of the scheme on energy prices remains limited compared to taxes).
[FI] Energy Efficiency Agreement for Industries	Monitoring and continuous evaluation done by Motiva Complementary external evaluation done in 2005	Monitoring and continuous evaluation reported to the Energy Authority and the Ministry of Economic Affairs and Employment, with a focus on evaluating the overall performance and relevance of the scheme, and looking for continuous improvements Monitoring and evaluation also used to report to the European Commission
[FI] Voluntary energy audits for municipalities	Monitoring and continuous evaluation done by Motiva, for the Energy Authority and the Ministry of Economic Affairs and Employment	Same as for the Energy Efficiency Agreement for Industries

Case	Stakeholders and roles	Evaluation purposes/expectations
[FR] Voluntary agreement for freight companies	Scheme monitored by the transport department of ADEME (French Agency for Environement and Energy Management) External evaluation commissioned and supervised by ADEME's evaluation department	Evaluation first done for internal needs (request from ADEME's transport department), with a synthesis published on ADEME's website. Evaluation criteria selected: • the intrinsic relevancy of the scheme (coherence between objectives & the sector needs and evolution) • the internal coherence (adequacy between the scheme and means) • the external coherence (adequacy between the scheme and other programmes) • the effectiveness (objectives reaching) • the efficiency (ratio cost-benefit) • impact measurement (unexpected effects)
[FR] "Future Investments" programme	Monitoring and ex-ante evaluation by ADEME of each candidate project External evaluation commissioned by ADEME Evaluation method discussed with the European Commission (to comply with the EU rules on State aids)	Reporting to CGI (public body in charge of supervising the programme) and to European Commission (compliance with EU rules on State aids) Assessing the direct and indirect impacts of the aid scheme (broad scope of indicators to capture the various effects of the programme) Analysing if the programme is adapted to participants' needs and is working cost-effectively + looking for opportunities of improvements
[DE] Energy Efficiency Networks Initiative	Scheme administered by the German Energy Agency (DENA) that registers the networks and verifies that they meet the minimum requirements. An independent institute (Fraunhofer ISI) monitors target achievements and evaluate the scheme.	Reporting obligations to the EU in EED Article 7 are in place.
[DE] Energy Efficiency Fund	Funds proceed from the Federal Ministry for Economic Affairs and Energy (BMWi) who also commissions the external evaluations. Schemes administered and monitored by different federal agencies or the development bank KfW.	Justification of the measures and use of public funding (cf. communication to Court of Auditors and Parliament). Identifying opportunities of improvements and getting new ideas for policy making. Need to define clear indicators to ensure consistency in the distinct evaluations of each of the measures included in the Fund.
[IE] Better Energy Homes	Monitoring and ex-post evaluation done by SEAI	Reporting to the Ministry of Finance, with a focus on impacts and cost-effectiveness of the scheme (CBA with different points of view)
[IT] White Certificates Scheme	General rules and targets set by the Ministries (Economy and Environment). Management and M&V done by GSE (public body in charge of stimulating energy services), with the technical support from ENEA (Italian Energy Agency) and RSE (technical centre owned by GSE). Penalties and DSO tariff allowance set by AEEGSI (Regulator of the energy markets).	At project level: validating the projects, verifying and crediting the amount of additional energy savings (in terms of white certificates). At scheme level: Monitoring the achievement of the targets and reviewing the corresponding costs (cf. cost recovery mechanisms). Analysing the market behaviours, and identifying whether rules need to be updated/adapted.

Case	Stakeholders and roles	Evaluation purposes/expectations
[LT] Renovation programmes with EU funding	The Ministry of Energy is responsible for the overall implementation, monitoring and verification of the energy efficiency policy. The State enterprise Energy Agency is responsible for estimating energy savings at the national level. The monitoring process involves persons receiving financial support from programs, the public authorities or bodies administrating programs implemented by the public authorities and the Ministry of Energy. In parallel, the National Audit Office did a review equivalent to a process evaluation.	Monitoring Rules defined by a national regulation set out the monitoring requirements to buildings, technological processes, installations or transport units receiving financial support from an EE policy measure. Main focus = monitoring the energy savings achieved. In parallel, the National Audit Report aimed at reviewing how the programme worked and how it could be improved.
[NL] Multi-year agreements in the industry	Monitoring done by RVO (Netherlands Enterprise Agency), based on the data reported by the companies committed to the voluntary agreements. External ex-post evaluations commissioned by RVO.	Reporting by RVO to inform the Energy Efficiency Consultative Group (OGE) and the Dutch House of Representatives about the progress of the scheme. Monitoring and evaluation are used to assess if the objectives defined by the participant companies/sector organisations are met, and to identify how to improve the scheme.
[NL] Subsidy scheme for housing corporations	Scheme set, implemented and monitored by the municipality of Amsterdam. Evaluation done by the Amsterdam audit office on their own initiative, with the support of energy experts from the Technical University of Delft about energy-related issues.	The Amsterdam audit office decided to evaluate this subsidy scheme because it is a practical example of the climate policies implemented by the municipality. One of the key objectives was to assess the actual impacts and cost-effectiveness of the scheme.
[NL] Fiscal incentives for cars	Tax measure under the responsibility of the Ministry of Finance Scheme administered by RVO (Netherlands Enterprise Agency) Ex-post evaluation done by PBL (Dutch environment agency) on their own initiative. PBL also carried out an ex-ante evaluation of the purchase tax measure. The Dutch Court of Auditors has also reported on the purchase tax measure.	Objectives of PBL: to get a better insight in the effectiveness, opportunities and barriers of different policy instruments, and more specifically to get a better picture of the effectiveness of the fiscal measure
[Nordic Countries] Nordsyn	The evaluation approach can be considered at two levels: At the European level, under the responsibility of the European Commission, for the evaluation of the overall impacts of the Ecodesign Directive and related regulations; At the national level, under the responsibility of each Member State, for the monitoring and evaluation of the implementation of these regulations, as market surveillance is under the responsibility of each Member State. Nordsyn is a cooperation programme between the Nordic market surveillance authorities. The Swedish Energy Agency, being the project manager, supervised the evaluation done by an external consultant.	One key objective of the evaluation was to assess the cost-effectiveness of market surveillance, by comparing the energy savings that would be lost without market surveillance and the costs of market surveillance. The results were used to get a political support for the cooperation programme to continue, and to communicate towards other countries to convince them of the importance and effectiveness of market surveillance.

Case	Stakeholders and roles	Evaluation purposes/expectations
[UK] Supplier Obligations	The Department for Business, Energy and Industrial Strategy (BEIS, formerly DECC) prepares the legislation setting the scheme, does internal reviews and commissions external evaluations and studies. The energy regulator, Ofgem, administers the scheme, and in particular the monitoring and verification system. Energy companies must report monthly to Ofgem which carries out audits of the claimed savings. Energy companies must also organise a verification of samples of actions by external auditors. Ofgem performs further controls. In parallel, the National Audit Office, an independent public sector auditing body, also made a review of the scheme.	Monitoring and verification done by Ofgem is primarily meant to monitor target achievement, and to fine-tune the implementation of the scheme. Internal reviews by BEIS and external evaluations and studies are done to update the deemed savings and identify needs for updates in the scheme settings. The review done by the National Audit Office aimed at evaluating the value for money of a range of energy efficiency schemes including the Supplier Obligation.
[UK] Warm Front	Scheme funded by the Government and administered by contractors also in charge of monitoring the scheme. The ministry commissioned a large research project at the beginning of the scheme, and several external evaluations along the scheme implementation. The ministry (DECC) also asked BRE (Building Research Establishment) to perform a technical evaluation. The scheme was also reviewed by the National Audit Office in 2003 and 2009.	Monitoring used to prepare annual reports to the ministry (with a focus on performance indicators). The research project was primarily focused on health impacts, but also looked at intermediate impacts such as changes in energy consumption. The external evaluations were mostly process evaluations (focused on how the scheme worked and how it could be improved, as well as on the targeting of the scheme to revise the eligibility conditions, etc.). The technical evaluation by BRE aimed at assessing the energy and CO_2 impacts of the scheme. The National Audit Office looked at both, cost-effectiveness (value for money) and process evaluation (see above).
[US] Auctions for capacity markets	Forward Capacity Market ran by ISO-NE (Independent System Operators of New England) that organises annual Forward Capacity Auctions. Providers of customer-based demand-side resources (e.g., utilities) can submit bids including detailed M&V plans that are then evaluated by ISO-NE or its contractors (external consultants).	The objective of the EM&V is to ensure that the capacity resources selected through the auctions are reliable and that the demand for electricity is and will be met, particularly during peak periods.
[US] Weatherization Assistance Program	DOE provides grants to states (Grantees), and states provide grants to local weatherization agencies (Subgrantees) to weatherize homes occupied by income-eligible households. Grantees need to report their activities and results to DOE that supervise the overall monitoring of the programme. ORNL (Oak Ridge National Laboratory) is in charge of the evaluation of the WAP at the national level. ORNL does its own evaluation and supervises also external evaluations commissioned by DOE. For these evaluations, ORNL prepares evaluation plans that are used for the tenders.	Primary evaluation objective: assessing if the programme is cost-effective. Evaluations are also used to see how the programme can be improved. Evaluations are communicated to the US Congress that uses them to see if this is a programme worth funding. As the main evaluation reports are public, they are also used by other stakeholders, such as NASCSP (National Association of State Community Services Programs).

3.2 Importance of legitimacy and credibility for stakeholders to have confidence in the evaluation results

Main conclusion from the feedbacks collected: stakeholders have confidence in the results if they trust the evaluators and their methods.

Some of the interviews also raised the issue of what "independent" can mean when speaking of "independent" evaluation.

Table 6. Feedbacks collected about the issues of confidence, legitimacy or credibility related to evaluation.

Case study	Quotes / comments collected about confidence, legitimacy, credibility, etc.
[AT] Environmental Support scheme	"Independence of evaluators is important but on the other hand evaluators have to know the scheme very well in order to understand the reasons for certain design choices policy makers made. This makes it sometimes difficult for authorities to find the right evaluators." The evaluation commissioner highlighted the importance of qualification and independency of the evaluator. The search of possible evaluators might be a difficult task, as wanted evaluators need to have a lot of knowledge about the scheme itself but cannot be associated with the Federal Ministry or one of the funded projects. Also the evaluation team needs to gather different skills (particularly technical and economics expertise) The independency and objectivity of the evaluator is taken very seriously in order to get credible results out of the evaluation. It is defined as independent from the body administrating the scheme, the Ministry commissioning the evaluation and project owners who applied for public aids.
[AT] City Energy Efficiency Programmes of Vienna	"Evaluation is time-consuming for both the evaluator and the evaluation customer. External evaluation is definitely helpful and recommended as it gives an outside view."
[BE] Primes Energie	"The issue of evaluators' independency should be considered in a pragmatic way. In the case of the Primes Energie scheme, the evaluation is done by a different service than the implementing service, but both services belong to the same department (Energy Department). On the one hand, one may say that this is not enough to ensure an independent evaluation. But on the other hand, this made that it was much easier for the evaluator to access the data needed for the evaluation, as programme managers may be reluctant to communicate data to persons outside their administration. They may have concerns about what the data will be used for, and about the way the evaluation conclusions will be drawn and communicated. Indeed, the adoption and ownership of the evaluation results require trust between programme managers and evaluators. So our experience is that the evaluation has more chances to be used, and recommendations to be implemented, when it is an internal evaluation."
[CR] Energy renovation of public sector buildings	Actual energy consumption is monitored through the Energy Management Information System (ISGE), managed by APN. If the energy savings are not actually realised, the ESCo does not receive the compensation from the public client (as per EPC).
[DK] EEO scheme	"The legitimacy of the evaluation is an important issue, so that the evaluation results can be used to take decisions for improving the scheme. When preparing the evaluations, there have been preliminary discussions with the energy distributors to discuss about the evaluation objectives and methodologies. We also created an international advisory group where evaluation experts provided comments and suggestions as external reviewers." "A point worth mentioning is that policy administrators should not rely on third party evaluations alone to tell them what is going on. Evaluations can never replace the trust and insights gained from a regular and frequent contact between the policy administrator and the affected parties."

Case study	Quotes / comments collected about confidence, legitimacy, credibility, etc.
[FI] Energy Efficiency Agreement for Industries	A well-functioning monitoring system for the agreement scheme has had a central role in revealing the results, create trust and credibility among all parties (not just the government) and in achieving long-term top-level commitment. "The absolutely critical starting points are reliability and coverage of data used in evaluation and the skills and technical know-how of the evaluators."
[FI] Voluntary energy audits for municipalities	"Comprehensive and reliable data was very important during the first years of the scheme for marketing purposes. If the audience is, e.g., managers of city hotels, then the results and experience presented must be from hotels located in city centers." "Double counting is not a problem when the monitoring boundaries are properly designed." "Concerning energy audits and voluntary agreements it is so normal thing to have that we don´t question it. It is a sort of guarantee that what we do has a clear and unquestioned justification."
[FR] Voluntary agreement for freight companies	The evaluation was requested by the programme manager, and supervised by ADEME.
[FR] "Future Investments" programme	Evaluation methodology to be approved by the European Commission (due to State aid rules). Use of an intermediate evaluation to test the feasibility of the evaluation.
[DE] Energy Efficiency Networks Initiative	"The monitoring institute does not get in direct contact with the network participants to guarantee the independent evaluation."
[DE] Energy Efficiency Fund	"One of the important features in policy evaluation is always the reputation of the evaluator. It determines the trust we can put into the results." "The important thing is, that they [the indicators and results] are well justified. For example in the exhaust heat evaluation, the survey sample size was too small to be used for a reliable calculation of free-rider effects."
[IE] Better Energy Homes	"The CBA was then complemented by the billing analysis to measure the impacts and verify the assumptions that were identified in the CBA as key sources of uncertainty (for ex., assumptions about comfort taking). This part of the ex-post evaluation made possible to have more robust results and to be able to justify them to the Ministry." "The quality of the evaluation work is key to bring credibility for the results." "One may have fear to do an ex-post impact evaluation, because it may show smaller results than based on the engineering estimates. However this increases the robustness of the results and therefore the confidence funders can have in them. (…) There were no more questions about the rationale or interest to implement this scheme. At the opposite, the questions were about how to make the scheme grow. (…) Empirical verifications represent a small budget compared to the whole budget of the scheme. Our experience with the ex-post impact evaluation is that it is really worth it"
[IT] White Certificates Scheme	The scheme includes detailed rules about what information should be submitted to GSE, how energy savings should be calculated, etc. There have been debates about the ex-post controls of the projects that could create financial risks (in particular for small ESCos).
[LT] Renovation programmes with EU funding	Investigation done by LEI on a sample of 80 buildings (2400 dwellings) raised the question of uncertainties related to scaled savings based on energy audits.
[NL] Subsidy scheme for housing corporations	The evaluation was done by the Amsterdam audit office on its own initiative. The audit office is an independent body. In the beginning of the evaluation process, the housing corporations did not give their cooperation.
[Nordic Countries] Nordsyn	"There are different discussions in different countries, but one is constant; money and budget. The question is always posed whether the surveillance is needed, and at what cost. The study should demonstrate to countries doing no surveillance at all that it pays off, and should motivate those already conducting some surveillance, to do more."

Case study	Quotes / comments collected about confidence, legitimacy, credibility, etc.
[UK] Warm Front	"Focusing evaluations on the "did it deliver" question leads, in terms of evaluation methodology, to a focus on sample size and representativeness. The main issue being to ensure that the method, and thereby the results, are robust enough."
[US] Auctions for capacity markets	"We need to ensure that resources can provide capacity when needed to meet projected resource adequacy needs - this requires a robust EM&V system to demonstrate how load reductions from EE will be quantified to meet capacity needs." For this reason, the monitoring and verification rules governing efficiency and demand response resources are very strict in capacity markets. The measured and verified electrical energy reductions during defined peak hours are the basis of payments and charges to market participants – in the case of under-delivery, participants would face penalty for the shortfall.
[US] Weatherization Assistance Program	"it was important to have an independent evaluation committee. It provided an external look that helped improve evaluation design and also brought a kind of validity stamp that the evaluation was done thoroughly and without bias. This gave legitimacy to the evaluation and helped getting the support from stakeholders."

3.3 Examples of good practices and difficulties to increase stakeholders' interest and confidence in evaluation

Table 7. Examples of good practices and difficulties to increase stakeholders' interest and confidence in evaluation.

Case study	Good practices / difficulties to increase stakeholders' interest and confidence in evaluation
[AT] Environmental Support scheme	Calculations submitted by applicants systematically checked when processing applications. Commissioning of an external evaluation every three years (with independence of the evaluators being one of the key criteria to select the evaluators) + involving an active steering group A plausibility check of the evaluation results was done by comparing them with national statistics and previous results. This ensures that the results presented are realistic.
[AT] City Energy Efficiency Programmes of Vienna	Monitoring and evaluation supervised by a coordination body that has exchanged with other City departments, the national monitoring body and the external evaluators to organise data collection and monitoring. Plausibility check by comparing top-down and bottom-up approaches. External ex-post evaluation planned every three years
[BE] Primes Energie	Internal evaluation but done by a service different from the one implementing the scheme. This ensures both, an external view and a good access to data and key contacts. The fact that the recommendations come from other public servants also increases the chances that they will be used.
[CR] Energy renovation of public sector buildings	"The next step in the development of this management tool is to connect the database of implemented projects [SMiV] with the consumption database [ISGE] which will then provide information on pre- and post-implementation energy consumption. This can currently be done manually, but the intention is to have this analysis done automatically (including calculation corrections such as climate correction, occupancy rate, etc.)."
[CR] Individual heat metering in multifamily buildings	Ex-post studies were done to assess in what conditions the new regulation would be cost-effective, based on metered energy consumption.

Case study	Good practices / difficulties to increase stakeholders' interest and confidence in evaluation
[DK] EEO scheme	Preliminary discussions with the obligated parties (energy distributors) when preparing the evaluations, especially about evaluation objectives and methodologies. During the evaluation, an international advisory group including evaluation experts does an external review of the evaluation methodology, results and conclusions. This increases the credibility of the evaluation. Regular contacts between the public authorities and the obligated parties are also essential to create a favourable environment for trust between parties.
[FI] Energy Efficiency Agreement for Industries	Quality of the monitoring system (including plausibility checks and verifications) to ensure that the data used are reliable. Providing training (about reporting requirements) and regular feedback. Importance and of the skills and technical know-how of the evaluators, for stakeholders to have confidence in their work.
[FI] Voluntary energy audits for municipalities	Training and qualification schemes for energy auditors, to ensure the quality of energy audits and reporting. Motiva makes plausibility check of the data reported by the auditors and the municipalities. And a sample of actions is verified each year.
[FR] Voluntary agreement for freight companies	Internal process to select what programmes should be evaluated, involving ADEME's top management in the decision. Evaluation plans prepared by ADEME's evaluation department, ensuring the quality of the specifications for the calls for tenders. Supervision of the evaluation with a steering committee.
[FR] "Future Investments" programme	Use of a mid-term evaluation to test and compare different evaluation methods.
[DE] Energy Efficiency Networks Initiative	Evaluation and implementation are strictly separated. While the implementation and rough assessment is performed by the German Energy Agency (DENA), the evaluation at the end of the predefined network runtime is performed by Fraunhofer ISI that is not involved in the implementation of the network initiative. The data collection procedure is also designed to guarantee data safety and confidentiality.
[DE] Energy Efficiency Fund	Attention paid to the proven expertise of the external evaluators.
[IE] Better Energy Homes	First cost-benefit analysis done using data available from the monitoring of the scheme. Then complementary ex-post evaluations with a billing analysis to measure the impacts and verify the assumptions that were identified in the CBA as key sources of uncertainty (for ex., assumptions about comfort taking). "This part of the ex-post evaluation made possible to have more robust results and to be able to justify them to the Ministry."
[IT] White Certificates Scheme	Information in the requests for white certificates are reviewed by GSE before white certificates are issued (after technical evaluation by ENEA and/or RSE when needed). GSE also randomly checks ex-post whether the implemented project complies with the approved project and conducts on-site inspections during the implementation or useful lifetime of the project. Since 2017, the evaluation of energy savings from standard projects must include measurements on a statistically representative sample. This was decided to ensure a more reliable evaluation of energy savings for standardised actions.
[LT] Renovation programmes with EU funding	Studies have been done to compare scaled energy savings (based on energy audits) with metered energy savings (based on energy bills corrected for normalisation factors). These studies concluded that further investigations would be needed due to the spread observed in the energy savings and related costs.
[NL] Subsidy scheme for housing corporations	Possibility for an independent public body (here the Amsterdam audit office) to do an evaluation on their own initiative.

Case study	Good practices / difficulties to increase stakeholders' interest and confidence in evaluation
[NL] Fiscal incentives for cars	Possibility for an independent public body (here PBL, the Dutch environment agency) to do an evaluation on their own initiative.
[Nordic Countries] Nordsyn	"We tried to underestimate rather than overestimate the savings, and the results were still very positive." This choice was made because it was not possible to assess the uncertainties, therefore using conservative assumptions was a way to provide results as reliable as possible.
[UK] Supplier Obligations	Reported carbon savings have included in-use factors since the inception of the Supplier Obligation in 1994 (expanding and refining the factors over time) to take into account the following points: • differences in performance in-situ compared to laboratory testing and imperfect installation • natural variations in the thermal performance of structural and fabric elements that cannot be fully determined by the assessment, e.g. the possibility that the un-insulated walls have different U-values than the standard assumptions and that U-value varies across different parts of the wall. • comfort taking (rebound effect) by the household, where some households may choose to heat their homes to a higher temperature • the household failing to operate the product/system effectively
[UK] Warm Front	The evaluation project started at the beginning of the scheme was planned over several years and involving research teams from universities. This offered unique conditions for high-quality investigations. The objective was to use the same quality standards as for health studies.
[US] Auctions for capacity markets	ISO-NE has an extensive M&V Manual that document what efficiency resource providers must do to demonstrate that their resources can reliably deliver the committed load savings at relevant system peak. This manual summarises the methods that can be used to document savings and are consistent with the energy efficiency industry's International Performance Measurement and Verification Protocols (IPMVP). In addition, the M&V manuals provide guidance on assumptions that can be used with regard to baseline efficiency, specify levels of statistical precision that studies of peak savings impacts must have, specify how recent any studies being relied upon must be, and address a variety of other M&V issues.
[US] Weatherization Assistance Program	An independent evaluation committee was established to make an external review of the evaluation, from its design stage. This helped improving the quality and credibility of the evaluation. The evaluation was done about a period with a large increase in programme's budget. A high budget ($20 million) was also dedicated to evaluation, to enable a good data coverage thanks to data collection on large enough and representative samples.

4 | Key prerequisite: monitoring and data collection

The main conclusions from this part are:

MESSAGE 04: Monitoring and data collection are essential for making any evaluation possible.

MESSAGE 05: Selecting the most relevant data to collect is a continuous process.

This is illustrated in the information collected below. First about how monitoring, data collection and evaluation are organised (section 4.1), reminding the no-brainer that it is essential to plan data collection when designing or adapting the policy measures. Second about feedback on how the data to be collected are selected and/or how data collection procedures are improved over time (section 4.2). Third about feedback on the difficulties encountered with data collection (section 4.3). And finally with examples showing good practices about practices for data collection and monitoring (section 4.4).

A special topical case study was also focused on the linkage between monitoring and evaluation.

4.1 How monitoring and data collection are organised

The main conclusion from the feedback collected is a no-brainer always essential to remind: it is always more difficult (and sometimes impossible) to collect data afterwards.

However, organising data collection upstream can be challenging. First because priority might be given to implementation, making that monitoring issues are dealt with later. Second because it is not always obvious to identify what data will be needed for further analysis (see next section).

Table 8. Feedbacks about organising monitoring, data collection and evaluation.

Case study	Examples / quotes about organising monitoring and data collection
[AT] Environmental Support scheme	***Monitoring***: Monitoring & verification of the data submitted by the applicants is done by Kommunalkredit Public Consulting that is in charge of the implementation of the scheme. The verification includes a focus on the additionality of the projects (based on predefined criteria). ***Data collection***: For monitoring, the main data collected are the data included in the applications. Review and evaluation of the scheme are based on data monitored on a regular basis. An important preliminary work is also done by the Federal Ministry to make all the relevant information about the scheme available to the evaluators. Complementary data collection is done by the external evaluators through interviews of stakeholders. ***Evaluation***: Regular analysis is performed by internal analysts of the Federal Ministry for Sustainability and Tourism and focuses on ecological, budgetary and economical effects (annual reviews). In addition, “the programme is evaluated every three years by external evaluators. The requirement for evaluation is set by law.” Evaluation commissioner pointed the importance of an exact and clear definition of the scope, purpose and questions of the evaluation. The external evaluation focuses on the same topics as the annual reviews, but also covers organizational effects. The scope of the evaluation is usually already set by the previous reports to ensure comparability; however, new developments are also taken into account. The scope of the evaluation also depends on the budget available for the evaluation. For example, for the latest evaluation done in 2013-2014, all the evaluation objectives initially considered would have required a budget three times higher than the one available. It was therefore necessary to prioritize the evaluation questions.

Case study	Examples / quotes about organising monitoring and data collection
[AT] City Energy Efficiency Programmes of Vienna	***Monitoring***: Monitoring and ex-post verification is done separately for each scheme of the umbrella programme. The level of ex-post verification is the highest for subsidy schemes. ***Data collection***: Each City department in charge of a scheme has to report the results every three years to a coordination body that centralises the data that are the starting point for the evaluation that complements these data with interviews. ***Evaluation***: The Vienna City Council decided when approving the SEP in 2006 that the SEP coordination unit will have to report to the City Council about the progress of SEP every three years, i.e. in 2009, 2012 and 2015. The evaluations done in 2012 and 2015 were commissioned to the same consortium of external evaluators. "Evaluation is time-consuming for both the evaluator and the evaluation customer. External evaluation is definitely helpful and recommended as it gives an outside view."
[BE] Primes Energie	***Monitoring***: Monitoring is done by the service in charge of the scheme. ***Data collection***: "The first point is a no-brainer, but always useful to remind: evaluation should be thought from the start, meaning when designing and starting each new policy. This is essential to organise the data collection and to ensure that the data needed for the evaluation will be available. This also helps to optimise costs for data collection." A database (Alfresco) is used to monitor the financial incentives, and merged with the other databases and data sources about government policy measures for energy efficiency in buildings. ***Evaluation***: The ex-post evaluation of energy savings is mainly based on the data collected through the monitoring system. Ex-ante evaluations have also been done to estimate future impacts according to different assumptions about trends in the number of actions implemented. Energy savings calculations are directly performed from the monitoring database, using the technical data collected for each action and harmonised values set at national level for the remaining parameters (baseline situation). "The evaluation method is based on the recommendations of the European Commission, as the evaluation results are used to report to the European Commission, initially within the ESD framework, and now within the EED framework. The way to apply the evaluation method was adapted to the existing data collection used to monitor the scheme."
[CR] Energy renovation of public sector buildings	***Monitoring***: Monitoring of the programme is done continuously, through the online partform SMiV (system for monitoring and verifying energy savings). The use of this tool has been made mandatory by law for all energy efficiency projects granted a public aid. SMiV has been managed and coordinated by the National energy efficiency coordination body, which has been integrated into the Ministry of Environment and Energy mid-2018. ***Data collection***: For each project receiving a public aid, an engineering simulation and an analysis of future energy savings is performed before the start of the project. Those data are entered in SMiV, which provides the main data for the evaluation of the scheme. Simple indicators, such as total energy and CO_2 savings, specific energy and CO_2 savings, the total amount of investments and grants provided, and cost of energy savings or CO_2 avoided can be seen with the SMiV application at any given moment for all types of measures or for different sectors. Actual energy consumption is not monitored by SMiV. The results in energy and CO_2 savings are based on deemed estimates, unless data specific to the energy savings projects are available. ***Evaluation***: Detailed evaluation, i.e. an analysis focusing on cost effectiveness, is performed as new plans are being developed. Therefore, at least a rudimentary cost-effectiveness evaluation is integral part of every new plan for the new period of the programme. "There is no formal procedure within the Programme that includes an in-depth ex-post evaluation. Main results for energy savings within the Programme can be observed through the monitoring and verification tool that is SMiV and through the internal reports made by the Environmental Protection and Energy Efficiency Fund (FZOEU), even if they do not bring conclusive results in ex-post evaluation terms, since the calculated ex-ante savings are the ones that are taken into consideration. Detailed analysis and evaluation of the programme, alongside with identification of key obstacles, and lessons learned is performed within development of new programmes and plans."

Case study	Examples / quotes about organising monitoring and data collection
[CR] Individual heat metering in multifamily buildings	***Monitoring***: The on-going monitoring and verification is conducted through the (web-based) System for Monitoring and Verifying Energy Savings (SMiV) (see other Croatian case above) ***Data collection***: See details about SMiV in the line above about the other Croatian case, as well as in the special topical case study on the linkage between monitoring and evaluation. ***Evaluation***: SMiV provides, among other indicators, results in terms of deemed savings based on values set in the legislation. Then two evaluation studies were made using billing analysis (method 2, metered savings), in order to assess the actual impacts of the actions on samples of buildings. A first evaluation study was done in 2016. This study was commissioned by Ministry of Economy to evaluate the applicability of the technology (individual heat cost allocators). A second study done in 2017 assessed the economic feasibility of implementation of heat allocators in multi-family buildings, and the conditions under which heat allocators would be cost-effective. This study was commissioned by Ministry of Environmental Protection and Energy.
[DK] EEO scheme	***Monitoring***: Annual reporting by the obligated parties that provide the data for the monitoring and verification done by the Danish Energy Agency. ***Data collection***: "Another well know lesson is that evaluation is easier when data are collected on a regular basis along the implementation of the scheme. This may be seen as a burden, in particular by the obligated parties. But this is essential to make monitoring and evaluation possible." "When the scheme started in 2006, it appeared too cumbersome to develop a centralised database to collect data from the obligated parties. The situation has changed now with the new possibilities offered by the development of Information and Communication Technologies. This could be a way to improve the data collection in the coming years." ***Evaluation***: External ex-post evaluation done at the end of each period, to verify the impacts, analyse cost-effectiveness of the strategies used by the obligated parties, and prepare the next period (e.g., to improve the rules). "Our experience is that when preparing a tender for an evaluation, the specifications for the evaluation should be focused on defining clear evaluation questions. The choice of the evaluation methods to answer these questions should be up to the bidders. This makes possible to compare offers with different methodologies." "The call for tenders for an evaluation has a major influence on what can be done in the evaluation. A good call for tenders can pave the way for a good evaluation and vice versa."
[FI] Energy Efficiency Agreement for Industries	***Monitoring***: Reporting obligations and the monitoring system were planned at the same time as the policy. This has helped the common problem that momentum is lost at least partly if a decision on monitoring is only done when the policies are already in the implementation phase. "Without robust monitoring data produced by our monitoring systems, the long-running scheme (since 1997) would not have continued this long, probably for just a few years." For more details, see also the special topical case study on the linkage between monitoring and evaluation. ***Data collection***: Motiva provides the participants with guidelines, training and support for them to report the data required. Motiva then verifies the data, and makes analysis on a regular basis (annual reports). A one-off large budget item was the construction of a new web-based monitoring database in 2007-2008. ***Evaluation***: Annual review and evaluation of the scheme is based on the data from the monitoring system, and done by Motiva. A complementary external evaluation was done in 2004-2005 in order to review the achievements over 1997-2005 and prepare the next period, getting feedback from the stakeholders and looking at how the scheme could be improved.

Case study	Examples / quotes about organising monitoring and data collection
[FI] Voluntary energy audits for municipalities	***Monitoring***: Monitoring was planned from the start of the policy measure. ***Data collection***: See line above about the case of the voluntary agreements ***Evaluation***: Annual review and evaluation of the scheme is based on the data from the monitoring system, and done by Motiva.
[FR] Voluntary agreement for freight companies	***Monitoring***: ADEME's transport department monitors the scheme. ***Data collection***: Monitoring is based on data reported annually by the participant companies through a web interface. The ex-post evaluation commissioned by ADEME in 2013 started by examining the data transmitted by companies involved in the scheme through the web tool of the programme, that are the data used for monitoring the scheme. This was complemented by an online survey of participants, as well as by a survey and interviews other stakeholders (particularly to prepare regional case studies, as the scheme is implemented at regional level). ***Evaluation***: After 8 years of implementation the programme manager (ADEME's transport department) issued a request to ADEME's evaluation committee for the programme to be evaluated. Evaluation requests are examined by ADEME's evaluation committee (including top-management) that then gave mandate to ADEME's evaluation department to include this evaluation in the multi-annual evaluation plan of ADEME, and to prepare an evaluation plan and a call for tenders.
[FR] "Future Investments" programme	***Monitoring***: ADEME monitors the programme for the funds dedicated to energy and environment projects. ***Data collection***: The monitoring is based on the data reported by the project holders receiving a public aid. Additional data collection (including surveys and interviews) is done for the ex-post evaluation. ***Evaluation***: ADEME is also in charge of organising and supervising evaluation of the programme (for the funds allocated to energy and environment projects). There are two levels of evaluation to consider: 1) On-going basis with annual reporting: The ex-ante evaluation of each project (technical, economical, financial & regulatory) carried out by experts from ADEME to size the financial support. 2) Multi-year basis: The ex-post programme evaluation is required by the CGI (the French governmental organization in charge of supervising the programme) and the European Commission to comply with the EU obligation for large state aid programme and with the contractual duty for ADEME and the State. The ex-post evaluation has thus to follow the requirements published by the European Commission.
[DE] Energy Efficiency Networks Initiative	***Monitoring***: Participants report to the scheme administrators (DENA) that does a rough assessment of future energy savings, based on the targets defined by each network. DENA also transfers anonymised data to an independent institute in charge of evaluating the scheme. ***Data collection***: "Data transmission channels are different between the networks and depend on the level of trust within the network. Some participating companies exchange savings and other data between each other, some only transmit them to the administrators. Some companies and networks include an energy consultant as an intermediary for the monitoring commitments." ***Evaluation***: After the end of the network operation time, a detailed evaluation is performed by an independent institute, using a survey to participating companies. Actual energy savings are collected and compared with the savings targets. For verification of the survey responses, a randomly selected sample of 10% of the companies is required to deliver savings documentation to the evaluators. However, no on-site physical measurements are performed by neither the scheme administrators nor the independent institute. In parallel, in the scope of the NAPE, a top-down estimation of savings has been performed.

Case study	Examples / quotes about organising monitoring and data collection
[DE] Energy Efficiency Fund	***Monitoring***: The Fund includes many different schemes. Each scheme is monitored and evaluated separately. Each project receiving a public aid from the Fund is reviewed ex-ante before approving the public aid. No ex-post verification is conducted. However, separate monitoring projects are implemented for certain measures of the Energy Efficiency Fund (EEF). ***Data collection***: In order to ensure consistency, indicators to be monitored have been defined in details, and a methodology report was elaborated to specify which data is necessary for an aggregated evaluation. This methodology is used by the different external evaluators who evaluate the different schemes. ***Evaluation***: The Fund is re-evaluated and updated each year until its end in 2020. The individual policy measures as well as the Fund as a whole are evaluated by independent entities regularly using both qualitative and quantitative approaches depending on the scheme.
[IE] Better Energy Homes	***Monitoring***: Regular monitoring of each application for a grant, together with random controls (including on-site inspections) ***Data collection***: Data collected through the monitoring of the applications include the amount of grants approved, number and type of actions carried out. This makes possible to calculate results in terms of energy savings, CO_2 emissions avoided and jobs supported. The ex-post studies did additional data collection (see details below). ***Evaluation***: Complementary ex-post studies: a survey of over 10000 participants (2010), a Cost-Benefit Analysis (2011) and an ex-post impact evaluation about actions implemented in 2009, including billing analysis (2012-2013). In 2017, a study reviewed the technologies supported by BEH to look at where Government support will be most effective. Results were expected by the end of 2017 with implementation of the recommendations from 2018.
[IT] White Certificates Scheme	***Monitoring***: GSE reviews the documentation of each application file and validates them (after technical evaluation by ENEA and/or RSE when needed) within 60 to 90 days. Then GME issues the certificates. ***Data collection***: Obligated or eligible parties (e.g., ESCos) can submit online application files. ***Evaluation***: Energy savings are first evaluated for each project submitted for white certificates, based on the rules defined for the scheme. Then GSE performs an annual review of the target achievements based on the reports of the obligated parties. GSE also randomly checks ex-post whether the implemented project complies with the approved project and conducts on-site inspections during the implementation or useful lifetime of the project.
[LT] Renovation programmes with EU funding	***Monitoring***: By law, each energy efficiency project receiving a public aid shall be monitored. The Housing Energy Efficiency Agency Agency evaluates and approves submitted investment plans and procurement documents. The Housing Energy Efficiency Agency then reports the data to the State enterprise Energy Agency. ***Data collection***: The public body in charge of implementing the policy measure collects the data from each project holder that received a public aid. For this scheme, project holders register the building or entity's indicators (heat consumption in kWh/m² per year, based on the building energy certificates) during the same calendar year and for one calendar year afterwards and then transmit the data collected to the Housing Energy Efficiency Agency that also collects the energy audits done when preparing the investment plans. The monitoring rules set by law also ensure that the following data are collected : type and number of actions implemented, energy properties of the actions, amount of investment for each action. ***Evaluation***: The State enterprise Energy Agency is responsible for estimating energy savings on the national level, based on the data reported by the Housing Energy Efficiency Agency.

Case study	Examples / quotes about organising monitoring and data collection
[NL] Multi-year agreements in the industry	***Monitoring***: The NL Agency is in charge of monitoring the scheme, based on the data reported by each company committed to a voluntary agreement. These data provide the basis for the sector reports that are discussed each year with the members of the Dutch energy-saving consultative body (OGE) of the sector. The monitoring reports submitted by the companies to the NL agency are checked for completeness and correctness by external consultants. This corresponds to a probability check, by comparing the development of energy use with production and project data. In addition, each company shall do an energy audit when committing to the scheme. These energy audits are reviewed by the NL Agency. ***Data collection***: Each company that signed the agreement is required to provide the NL Agency with monitoring data before 1st April each year. Data requirements are focused on progress made by the companies with implementing their action plans and their practice of systematic energy management. ***Evaluation***: For each period of the agreement, the NL Agency also commissions an external ex-post evaluation.
[NL] Subsidy scheme for housing corporations	***Monitoring***: The scheme is monitored by the municipality of Amsterdam based on the data reported by the housing corporations. ***Data collection***: The data reported by the housing corporations include the energy label of buildings before and after renovation (number of label steps). For the ex-post evaluation, other data sources were needed to compare energy consumption estimated by the energy certificates and actual energy consumption (from energy bills). ***Evaluation***: The Amsterdam audit office performed an ex-post evaluation on its own initiative with the technical support from the Technical University of Delft.
[NL] Fiscal incentives for cars	***Monitoring***: The tax measure is monitored by RVO and the Ministry of Finance. ***Data collection***: Monitoring is focused on the data on car sales (and corresponding energy labels). PBL (Dutch environment agency) used the monitored data on the sales of new cars (available in the national statistics) as input data for a stock modelling (based on economic modelling). ***Evaluation***: PBL did an ex-post evaluation on its own initiative.
[Nordic Countries] Nordsyn	***Monitoring***: The national market surveillance authorities (MSA) are in charge of monitoring the implementation of the requirements of the EU EcoDesign and energy labelling directives. The Swedish Energy Agency coordinates the Nordsyn project, a cooperation project between MSA of Nordic countries. ***Data collection***: The Swedish Energy Agency worked with the other MSA to gather the best data available from each country (mostly from laboratory tests on samples of appliances). ***Evaluation***: An external consultant was contracted to evaluate Nordsyn. This study was the Effect project. "The [Effect] project took at least a year, not full-time work, but to be able to assemble the data. Again, the NORDSYN and its synergies proved crucial for collecting data for the Effect study."
[UK] Supplier Obligations	***Monitoring***: The energy regulator Ofgem is responsible for monitoring and verifying the reported energy efficiency actions. Energy companies report monthly to Ofgem about the actions implemented and corresponding carbon savings achieved (according to deemed scores per type of action). They also have to organise verifications on samples (5%) of the actions by external auditors. Ofgem reviews the data reported, and performs further controls. It also monitor target achievements. ***Data collection***: Monitoring is based on the data reported by the energy companies, complemented by further controls. BEIS uses other data sources, and particularly NEED (National Energy Efficiency Data-framework) that includes for example data from energy suppliers on metered energy consumption.

Case study	Examples / quotes about organising monitoring and data collection
	Evaluations are doing further data collection (e.g., surveys of participants, interviews with stakeholders). ***Evaluation***: In the UK every new public policy, including a new phase of the Supplier Obligation, is subject to an impact assessment prior to its implementation (usually compiled by the relevant government department) that can be classified as an ex-ante evaluation. The ministry in charge of energy (now BEIS, formerly DECC) carries out interim reviews of the Supplier Obligation based on Ofgem's and their own analysis of the scheme. Those reviews are then used to design the subsequent phase of the Supplier Obligation. In particular, BEIS continually refines the energy saving estimates of specific interventions, which are used for monitoring the scheme and to calculate the overall impact of the Supplier Obligation. This is done by using data from NEED (National Energy Efficiency Data-framework). The ministry (BEIS and formerly DECC) has also commissioned independent ex-post evaluations of the Supplier Obligation. "The Supplier Obligation is continuous and there is typically not enough time to carry out an evaluation at the end and apply the learnings in the next phase. We therefore conduct post implementation reviews which happen mid-term during an ongoing phase of the Supplier Obligation. There is also continuous tracking of impacts of scheme in terms of actions and costs - this reduces the need for substantial evaluations at the end of the scheme." "Evaluations only form subset of how BEIS reviews policies. The science team at BEIS commissions studies on new technologies, costs and barriers. We also talk to stakeholders involved in delivery to understand their perspective. For example, we returned to deemed savings as result of stakeholder feedback."
[UK] Warm Front	***Monitoring***: The ministry contracted scheme administrators that were in charge of an on-going monitoring with an annual synthesis to be reported to the ministry. ***Data collection***: The contracted scheme administrators collected data from the participants and other stakeholders involved in the implementation of the scheme. More specifically, quality assurance processes included post-installation inspections about the quality of the actions. They did not included ex-post verifications of energy savings. Energy savings were monitored based on simplified engineering calculations (using building energy rating). The research and evaluations did further data collection (see details below). ***Evaluation***: The ministry commissioned several external studies depending on its evaluation needs. A study of changes in energy consumption for a sample of participants was made within the broader evaluation of health impacts in the first years of the scheme. Some years after, the ministry (DECC - Department of Energy & Climate Change) asked BRE (Building Research Establishment) to make a study where a modelling more sophisticated than the building energy rating was used, including a reduction factor to take into account direct rebound effect and possible underperformance of insulation actions. One suggestion made by one of the evaluators of the scheme in the interview for the EPATEE case study was that it would be interesting "to consider another evaluation approach, using the methodology of action-based research. This is closely linked to clarifying evaluation objectives and timing: is it to decide about future funding of the policy? or is it to identify how to improve the policy? In the first case, evaluation can be done on a period with no major change to the policy, with the main objective to observe its impacts. In the latter case, the earlier sources of improvements are identified, the better. But then changes to the policy may be done frequently, making it difficult to assess the impacts due to the decisions taken. Moreover, the first case cannot be used when working on time-limited problems such as climate change. Targets are high, and timelines are tight. We cannot afford to monitor long time series on large samples like usual academic criteria would require. We need to explore other ways to collect feedbacks, to be able to improve policies on an almost on-going basis." The National Audit Office also performed two reviews of the scheme (2003 and 2009) and reported to the ministry and the House of Commons.

Case study	Examples / quotes about organising monitoring and data collection
[US] Auctions for capacity markets	***Monitoring***: ISO-NE (Independent System Operators of New England) reviews the M&V plans provided by bidder applicants as part of qualifications packages to ensure that they comply with their M&V Manuals. ***Data collection***: Data are submitted along the annual auction process, and then once the projects are implemented, on an annual basis. ***Evaluation***: Once resources clear the market and are being delivered, ISO-NE reviews documentation from the project sponsors to ensure the reported load savings are consistent with the M&V plan and any M&V studies identified by the plan. Ex post peak savings verification entails conducting independent evaluation based on M&V plans and ISO-NE protocols. The outcomes are realisation rates, i.e. the difference between capacity obligation and verified capacity value, for various project/action categories, which are applied to the estimated peak savings of the EE resources for reporting to ISO-NE. In the past, expert M&V consultants have been hired to assist with this work. ISO-NE reserves the right to audit the load savings databases and related documentation of the efficiency project sponsors. "Before EE can participate in the capacity auction, a detailed M&V plan needs to be submitted. If qualified, a project enters the capacity auction. If it clears the auction, M&V is used to determine whether the project delivered the savings for which it received payments during the timeframes specified (peak period savings separated from non-peak period savings). Most projects are being evaluated as part of utility Energy Efficiency Obligations evaluations as they receive funding from both Energy Efficiency Obligations and the capacity auction - those utility evaluations also inform the M&V plans submitted to ISO-NE."
[US] Weatherization Assistance Program	***Monitoring***: "WAP has a strong monitoring component. DOE-WAP supervises the overall monitoring, gathering data from all states. Then each state monitors its Sub-Grantees. Quality insurance is thus implemented at different levels." Grantees (the states) are required to report quarterly to DOE. They are also required to conduct comprehensive monitoring of each sub-grantee (local implementing agencies) at least once a year. DOE performs weekly, monthly, and quarterly desktop reviews as well as site visits on grantees' performance. DOE overview also entails quality assurance (QA) visits. These quality assurance visits occur at the Subgrantee level. All the monitoring activities are summarized into an internal monitoring report for consideration and review during annual planning. ***Data collection***: Data reported by the grantees include data on their expenses, number of homes weatherized (quarterly reports) and on other performance data (annual reports). This is favoured by the use of a web-based interface, where state agencies can directly enter their data and also receive technical support (FAQ, etc.). The comprehensive monitoring of sub-grantees must include reviews of client files and SubGrantees records, as well as inspection of at least 5 percent of the weatherized units. ***Evaluation***: "There were first small evaluations, and then an in-depth and thorough evaluation of the 1989 program. Then evaluations were done at state's level with periodic meta-evaluation by ORNL to put data together and update the results from the 1989 evaluation. More recently, new national evaluations were done in 2008 and 2011, in connection with the increase in the budget due to ARRA (American Reinvestment and Recovery Act)."

4.2 How the data to be collected are selected

The main conclusion from this section is that identifying the most relevant data to be collected is a continuous process. Moreover, the experience from the various case studies also show that clear information, or even training, of the stakeholders about the data they are required to report is as important as the tools (reporting templates, spreadsheets or online platform) set up to collect the data.

Table 9. Feedback about how data are selected or how data collection is improved over time.

Case study	Examples / quotes about how data are selected or how data collection is improved
[AT] Environmental Support scheme	Management aspects (including application process and data requirements) are analysed every three years in the external evaluation + regular feedbacks between the implementation body (Kommunalkredit Public Consulting) and the Federal ministry More standardized procedures (e.g. online application) used from 2011, which helped reduced processing time by 39 days (20%) Emphasis in the monitoring on additionality criteria + evaluation of the average time to process applications
[AT] City Energy Efficiency Programmes of Vienna	When setting the initial monitoring and evaluation framework in 2009, there were several coordination meetings of the SEP coordination body with the national monitoring body and the external evaluators. This approach has been used to continuously improve the quality of monitoring. For all three evaluation reports (2009, 2012, 2015) the external evaluators and the coordination body developed standardised templates to report actions and other data on the instruments that were sent to the respective departments. The templates have been set taking into account exchanges with the national monitoring body. For the third report this process was complemented by interviews with persons in charge of the different schemes to sort out open questions and verify the submitted information.
[BE] Primes Energie	The review by the Court of Auditors in 2009 included an analysis of the scheme management that led to improvements and simplification in the way the applications were monitored and controlled by the administration. About the evaluation of the results, the Court of Auditors concluded that the database used by the administration to monitor the scheme was not registering all the information from the applications that would be needed to assess energy savings. This was further improved.
[CR] Energy renovation of public sector buildings	It is planned to connect the databases monitoring energy consumption in buildings and actions receiving public grants to enable more systematic ex-post verifications of energy savings.
[CR] Individual heat metering in multifamily buildings	Feedback from a participant to the scheme: “The project documentation was submitted, the financing was received with the very little hassle and the project moved on with no particular obstacles on the administrative side. There were no additional requests for reporting on actually achieved energy savings, which would not be a problem as these data are easily obtained.”
[DK] EEO scheme	The review done by the National Audit Office in 2017 concluded that the efforts made by the public authority to ensure that the scheme promotes cost-effective energy savings have been satisfactory, but that the efforts to ensure that the energy distributors comply with the scheme’s rules were not sufficient. The statutory auditors pointed that the annual sample checks done by the public authority covered a small proportion of the energy distributors’ reports, whereas high error rates were identified from 2013 (27%) and increasing to 43% in 2015. These problems were identified based on the regular monitoring of the scheme and the results of the evaluation done in 2015. The observations from the monitoring and evaluation of the scheme were thus already taken into account in the new agreement entered into force in December 2016. This agreement includes a greater focus on the control and documentation of the energy savings by the energy distributors. It was complemented by a decision of the government in April 2017 to allocate additional resources (€15 million, for 2017-2021) to the Danish Energy Agency and the Danish Energy Authority to strengthen the control on reported energy savings. This includes expanded random checks, special controls and enhanced control of the costs incurred by the energy distributors, thanks to the recruitment of about 20 new employees.
[FI] Energy Efficiency Agreement for Industries	“Data coming from the participating companies is the ’raw material’ of evaluation. There is no possibility to make compromises in the quantity, quality or submission deadlines of this data. Ensuring timely submission of good quality data requires a lot of administrative work. This has involved a lot of discussions with the contact persons of participating companies and looking after. When the new contract for the period 2017-2025 was being negotiated with the sector, annual reporting was one of the few topics which were not open for negotiation – it was a must.” The participants have a possibility to report on their overall satisfaction to the agreement scheme in their annual reporting. Additional feedback is collected in various events organized around the

Case study	Examples / quotes about how data are selected or how data collection is improved
	agreement scheme. Feedback was also the topic of the formal third-party evaluation carried out in 2005.
[FI] Voluntary energy audits for municipalities	See the line above about the experience of the voluntary agreements.
[FR] Voluntary agreement for freight companies	The data to be reported by the participants were selected to enable to inform the indicators used to monitor the scheme. The commitment to report these data is the counterpart for participants to receive technical support from the scheme and to be registered as "committed company". The data collected for the evaluation were selected in order to assess the selected evaluation criteria and taking into account feasibility.
[FR] "Future Investments" programme	An intermediate evaluation has been done to test the evaluation methodology (including data collection and processing).
[DE] Energy Efficiency Networks Initiative	"From an efficiency and effectiveness point of view, the evaluation practice is very good. It requires relatively little effort. However, with a larger base of data more detailed analyses would be possible. These would make it possible to supply participants with benchmarks that could help them to improve the network performance. However, in an inquiry among participants whether they are interested in further data, only about half of the present company representatives showed interest. The whole monitoring appears to be geared towards simplicity. The survey questionnaire has been shortened because company representatives recommend imposing as little effort as possible to participants."
[DE] Energy Efficiency Fund	In order to aggregate the results from all the quantifiable measures included in the Energy Efficiency Fund, a structured spreadsheet template was elaborated that automatically calculates the aggregated values for all indicators. Having an identical spreadsheet with all values allows combining these sheets in an aggregation spreadsheet without the need to input the data for each scheme individually. Tables with the exact number of rows and columns from the spreadsheet are then included in the template for the report, which makes it easy to represent the data for all policy measures in a uniform way.
[IE] Better Energy Homes	An important source of uncertainty about the CBA was the share of costs paid by the participants that needed to be assessed, as it was not directly monitored at that time (2011). Monitoring has now been changed to record both, the amount of grants paid and the total investments made (including the amount of grants + the amount paid by the owners, possibly with other financial support like from energy companies due to the energy efficiency obligation scheme).
[IT] White Certificates Scheme	The Ministries (Economy and Environment) set the rules, including the data to be included in the requests for white certificates, as well as in the annual reports of the obligated parties. These rules have been fine-tuned over time, taking into account feedback from GSE as well as from stakeholders.
[LT] Renovation programmes with EU funding	Monitoring Rules set out the monitoring requirements to buildings, technological processes, installations or transport units receiving financial support from energy efficiency programmes implemented by the public authorities. The monitoring is made to be relevant to programme administrators who evaluate individual indicators, review the monitoring exercise, and report to the Ministry of Energy.
[NL] Multi-year agreements in the industry	"Since the government aims to minimise the reporting burden on companies, reporting requirements are kept simple." In January each year, companies that take part in the monitoring receive a letter containing instructions on how to submit the reporting information. The information that must be completed for the annual monitoring relates in particular to the energy consumption and the energy-saving measures that were implemented in that year.
[NL] Subsidy scheme for housing	During the roll-out of the subsidy scheme the municipality of Amsterdam only controlled the energy label of the houses after renovation. These data were not sufficient for the evaluation. The evaluation could finally be done by combining databases from several institutions: data about

Case study	Examples / quotes about how data are selected or how data collection is improved
corporations	energy use from CBS and energy labels of the energy labelling registration database from RVO.
[NL] Fiscal incentives for cars	The ex-post evaluation as carried out by PBL is primarily based on national statistics about the sales and stock of cars. These data are the key inputs for DYNAMO, the model used to assess the effects of the tax measures on the sales and stock of cars.
[Nordic Countries] Nordsyn	"There are many ways to conduct such a study and obtain data. There is always a cost and time needed to obtain data, and there were of course many estimates used in the study. It could have been conducted in many ways, but the idea was to prove that savings are occurring and what can be saved, as well as how cost-efficient the market surveillance is."
[UK] Supplier Obligations	Ofgem defines what data the energy companies must report, among other scheme rules, after an open consultation of stakeholders (particularly energy suppliers). In parallel, the ministry (DECC then BEIS) developed the National Energy Efficiency Data-framework (NEED) as the continuation of data collection work started in 2004 making more disaggregated data available for the analyses done by the Ministry (especially data from the energy suppliers about metered energy consumption).
[UK] Warm Front	The data collected from the participants and other stakeholders involved in the implementation of the scheme were defined by the scheme administrators in order to verify that requirements were met (e.g., eligibility) and to be able to monitor the scheme and report about the performance indicators included in the contract with the ministry. The data collected for the external studies were selected according to the evaluation objectives and the evaluation methods chosen by the evaluators. "One important lesson from the research project was that it could be more relevant to use smaller sample, but to collect more robust data and going more into the details. This makes possible to get a better understanding of how the policy works and why results are achieved or not. Which is often more valuable to policy officers than getting only a view on what was achieved."
[US] Auctions for capacity markets	ISO-New England (ISO-NE) requires bidders to deliver a detailed M&V plan as part of the prequalification process. The plan needs to contain a description of the equipment or types of equipment for projects being installed and/or modified, as well as of the approach taken to monitoring and verification. Data requirements are detailed in the M&V Manual prepared by ISO-NE.
[US] Weatherization Assistance Program	The monitoring is done with a bottom-approach: sub-grantees report to grantees (states) that report to DOE. Information to be reported are defined by DOE to inform the performance indicators. "The evaluations done in 2008 and 2011 represented an about $20 million effort, which made possible to have a dedicated data collection. One of the objectives was to verify the reliability of the monitoring. That's why it could not be based on monitoring data. Moreover, these evaluations went beyond the assessment of energy savings and costs, and thus required additional data collection, particularly for the part of process evaluation)" ORNL organized a National Weatherization Network Committee (i.e. weatherization officials, local weatherization officials, DOE staff, ORNL staff, and independent evaluators) to provide input and data for the ex-post evaluations.

4.3 Difficulties encountered with data collection

Table 10. Feedback about difficulties encountered with data collection.

Case study	Examples and quotes about difficulties encountered with data collection
[AT] Environmental Support scheme	"Data collection is a main issue. The data requested should be clear before evaluation starts because it is often impossible to gather data afterwards"
[AT] City Energy Efficiency Programmes of Vienna	Despite the intensive efforts of the coordination body and most City departments to continuously improve the data collection and document the results on a regular basis, there was still a lack of data availability to quantify the energy savings of some schemes on a reliable basis. This applied for example for schemes targeted at private service buildings. Another issue pointed by the evaluators was the reporting frequency (every three years), that

Case study	Examples and quotes about difficulties encountered with data collection
	was not harmonized with other reporting obligations for other City programmes or to Federal state. Which created additional work for several City departments. However it is also very likely that without the intermediate reporting, the final evaluation would have been much more difficult to perform.
[BE] Primes Energie	"The biggest difficulty is to justify the efforts (time and budget) needed to collect the energy-related data. This is particularly true for the evaluation of measures that have multiple objectives and/or for which energy efficiency is not the priority objective (for ex., the other scheme providing grants for dwelling renovation and aiming at eradicating substandard housing). The difficulty is then to find the right compromise between minimizing data collection efforts and ensuring the reliability of the results." "Another difficulty is the relation between "providers" and "customers" of data or databases. As mentioned earlier, data providers (here programme managers) may not fully trust data customers (here evaluators), and then not give easily an access to the data." "The efforts needed to collect data and perform the evaluation are really worth it. It is very useful to be able to assess the results at a given time horizon. This was the approach of the ESD (reviewing the energy savings achieved in 2010 and then in 2016). And this is in line with the governance of the AEE for example. However, we do not see the added value of the additional efforts of data processing due to the requirement of the EED article 7 to assess the energy savings over a given period (2014-2020) that is not in line with the way the actions are monitored for the scheme management. And neither is it consistent with the temporality of the impacts of the actions implemented to achieve the overall 2020 target." "The changes in the application process were mostly due to the requirements of the EED (for ex. about materiality and additionality). This shows how the rules applied to monitor and evaluate the energy savings may sometimes affect the way the measure works, and even its effects."
[CR] Energy renovation of public sector buildings	Initially, no connection between databases monitoring energy consumption and actions respectively; this was under consideration and could drastically improve calculation's accuracy.
[DK] EEO scheme	"Another issue that appeared difficult to handle is the verification of the situation before the implementation of the action. Most of the M&V rules are used to verify the actions ex-post. However, it is often not possible to check the "before" situation once the action is implemented. For example, it is difficult to verify what boiler (type and age) was replaced or what level of insulation was already in place. One way can be to ask for a picture of the building or equipment before the implementation of the action. However, it is still possible for installers to "cheat", for example by photographing another boiler. So this remains a concrete tricky issue." "The main constraints were as often is the case related to time and budget limits. The main challenge in the 2008 and 2012 evaluations was to deliver sound results fast enough, with a very tight schedule. This implies finding ways to prioritize and select what work to do – in particular, in relation to the data collection." "One thing that we experienced, for example in the 2012 evaluation, was poor quality of the data reported by the obligated parties to the Danish Energy Agency. Each obligated party had its own information system to gather and report data. Some of the data was not accessible electronically. This means that data are available in different formats, making it difficult to put them altogether in a consistent way and also to verify them. The Danish Energy Agency was in favour of a common platform to centralize and report data from the start of the scheme in 2006. But the stakeholders were strongly opposed to that. These positions have since changed and it may in future evaluations be possible to access data from a common database." Difficulties were also encountered to get market data that would enable a quantitative assessment of the impacts on EE markets. Likewise, there were difficulties to get disaggregated-enough data and consistent long-time series to perform econometric analyses per end-use sector. As the costs reported by the energy distributors increased significantly (by about 25%) between 2010 and 2016, the National Audit Office was also asked to review the scheme, including a review of the written materials about the scheme, interviews with key stakeholders and an audit of a sample of energy distributors. The statutory auditors concluded that the efforts to ensure that the energy distributors comply with the scheme's rules were not sufficient. Annual sample checks

Case study	Examples and quotes about difficulties encountered with data collection
	done by the public authority were estimated to cover a too small proportion of the energy distributors' reports, whereas high error rates were identified. There was a lack of verification if the energy distributors had corrected the energy savings where errors were identified, as well as not enough control of the risk that energy distributors bought energy savings to their affiliated companies at an overcharged price. More generally the auditors warned that the costs reported by the energy distributors were not enough monitored. All these issues have been tackled in the new agreement entered into force in December 2016.
[FI] Energy Efficiency Agreement for Industries	Investment cost and time needed to fine-tune the web platform
[FI] Voluntary energy audits for municipalities	See the line above about the case of the voluntary agreements
[FR] Voluntary agreement for freight companies	The evaluation was limited by time and budget: it prevented evaluators to get a representative pool of not committed transport operators (in order to better understand why they are not involved in the programme).
[FR] "Future Investments" programme	The complexity of the project and the fact that R&D projects might be adapted over time (compared to the initial plans submitted to get a public aid) lead to possible errors in the monitoring database (difficulties to ensure systematic updates). Energy savings cannot be monitored or verified ex-post on short term (as they will occur after a certain time depending on technology development and time to market).
[DE] Energy Efficiency Networks Initiative	"The monitoring is running since autumn 2017. The first reliable data for network performance will hence only be available towards the end of the network action period [2020]. Until now, the sample size for evaluation is small, but it will be larger in the next evaluation round."
[IE] Better Energy Homes	"We faced some difficulties for the billing study to access billing data. The main issue was that for legal reasons, we had to contact participants to get their agreement to use their billing data for the evaluation (even if these data were anonymised). This took a lot of efforts and led to a smaller sample compared to the initial plan. We therefore strongly recommend to put a condition when setting the scheme to ensure the access to data, and that the legal team checks the validity of this condition. This should be possible for any grant scheme, as participants will very likely agree on this condition to be able to get the grant. This can save a lot of data collection efforts." Also difficulties in handling very large sets of data, which may be technically difficult and requires to be very cautious for respecting the privacy rules about data.
[IT] White Certificates Scheme	Despite the systematic review of all files submitted for white certificates requests and ex-post controls, limiting errors and frauds is an on-going issue. Moreover, the strengthening of the rules to monitor energy savings from specific projects (e.g., large projects in industry) has increased the costs of these projects.
[NL] Multi-year agreements in the industry	"This [the fact that reporting requirements are kept simple] has implications for the reliability of the data available from the monitoring of the scheme, e.g. for the 2008-2012 ex-post evaluation. It is difficult to cross-check the energy savings claims made by the companies and sectors as a whole." Further data collection (interviews of stakeholders and survey of participants) was needed when performing the ex-post evaluation.
[NL] Subsidy scheme for housing corporations	It was not easy for the evaluators to follow the history of the policy measure. People that know the history of the policy measure, the implementation details, etc. are difficult to find in a large organisation. About the data monitored by the municipality of Amsterdam, the evaluator found a lot of missing data and mistakes and concluded that there were not enough controls and checks in the execution of the subsidy scheme. Linking the databases that could be used for the evaluation was costly and time consuming. For example, the energy labelling registration database included often incorrect or inaccurate information about the situation in practise, so the reliability of registered energy labels was

Case study	Examples and quotes about difficulties encountered with data collection
	weak.
[NL] Fiscal incentives for cars	The impact evaluation done by PBL was focused on the reduction in CO_2 emissions, through the indicator of specific emissions of new cars (in gCO_2/km). Several major assumptions were thus needed to assess energy savings from this indicator.
[Nordic Countries] Nordsyn	The approach chosen was to gather the best data available, and then to complement by using conservative assumptions. Several major assumptions were thus needed, creating uncertainties. "Providing a larger data input would surely produce even more accurate results. Also, in our study we used actual sales data only for Sweden, and we used the Danish model to estimate the sales data for the rest of the countries. We would definitely recommend using actual sales data for as many countries as is available and possible."
[UK] Warm Front	"It is not rare to see large data collection ending in providing data that cannot be used for rigorous data analysis, due to various problems in the measurement and data collection chain. Such risks should not be neglected."
[US] Auctions for capacity markets	"While the output of generators are directly measured with meters and outage rates can be forecast, reductions in load produced by energy efficiency is not as directly observable, which poses a challenge when considering energy efficiency for meeting resource adequacy."
[US] Weatherization Assistance Program	Reportedly the National Weatherization Network Committee faced a major challenge with the collection of utility bills. Main hurdles related to utility bill waivers that utilities require prior to releasing utility bill information (e.g. Subgrantees forget to ask their weatherization clients to sign waivers, utilities reject the standard waivers used by Subgrantees). "Future evaluations could be done more easily if the data needed would be collected routinely. For example, collecting utility bills afterwards is very difficult and time-consuming. Provisions should be taken at least to include utility waivers that participants would sign when getting WAP benefits." In the future, WAP managers might consider working with utilities to develop standards for utility waivers that the programme could then adopt for next evaluations. One challenge about this is that each state has its own privacy rules and its own public utility commission. Therefore, what is acceptable in one state may not be acceptable in another.

4.4 Examples of good practices about data collection and monitoring

Table 11. Examples of good practices about data collection and monitoring.

Case study	Good practices about data collection and monitoring
[AT] Environmental Support scheme	Regular analysis to identify changes, new trends, problems and needs to adapt the scheme requirements (e.g., eligibility criteria). More standardized procedures (e.g. online application) used from 2011. Complementary external evaluation to investigate how the procedures can be improved. To ensure that the evaluator will have a good knowledge of the scheme, the Federal Ministry has to gather all the information spread in various sources and over time. Checking and sorting the information is often needed to make them clear and usable by the evaluator. This should often be complemented by discussions with the evaluator along the evaluation, when further clarifications are needed. This work on the side of the evaluation customer can be time-consuming. But experience shows that it is essential for evaluation conclusions to be consistent with the actual implementation of the scheme (and not disconnected from the ground). Moreover, this also provides policy officers with a better understanding of the programme. Having a solid database was also pointed as a crucial point to enable a good evaluation.
[AT] City Energy Efficiency Programmes of	To achieve a consistent approach along departments a coordination body was installed that overlooks progress of the programme and is in charge of evaluation and reporting the results. The external evaluators emphasised that the coordination body has a very good overview of the status of

Case study	Good practices about data collection and monitoring
Vienna	SEP implementation due to the regular contacts with the relevant services. Some of the recommendations made by the evaluators dealt with monitoring and evaluation: defining as far as possible impact targets and indicators (in the sense of specifically formulated outputs and outcomes) that could serve as basis for the monitoring and evaluation; to provide the City departments in charge of the measures with tangible monitoring parameters (e.g., specifying energy indicators or other metrics) for data collection, which can guide or standardize the collection and calculation of savings and directly monitor the implementation of the measures.
[BE] Primes Energie	A database (Alfresco) is used to monitor the financial incentives, and merged with the other databases and data sources about government policy measures for energy efficiency in buildings. The combination of sources help to select the best data available and avoid double counting. Data collected include the invoices submitted by the applicants (making possible a verification).
[CR] Energy renovation of public sector buildings	Monitoring and verification conducted through the (web based) System for Monitoring and Verifying Energy Savings (SMiV), a tool defined by the legislature as obligatory. Any actor receiving public subsidies for an EE action/project must report the action/project on the online platform. "The next step in the development of this management tool is to connect the database of implemented projects with the consumption database which will then provide information on pre- and post-implementation energy consumption. This can currently be done manually, but the intention is to have this analysis done automatically (including calculation corrections such as climate correction, occupancy rate, etc.)." Actual energy consumption is monitored through the Energy Management Information System (ISGE), managed by APN. If the energy savings are not actually realised, the ESCo does not receive the compensation from the public client (as per EPC).
[CR] Individual heat metering in multifamily buildings	Same platform (SMiV) and requirements (see line above). Feedback from a participant: "As the representative of dwelling owners, I am monitoring these effects through data about the heat consumption and related bills. All tenants are informed about the effects through our building Facebook profile. This way, the awareness on energy consumption and consequences of our investment activities, but also our behaviour as energy consumers, is raised."
[DK] EEO scheme	The ex-post evaluations and review by the Court of Auditors have led to recommendations to improve monitoring and verification processes. This also helped to get additional funding for the monitoring and verification activities.
[FI] Energy Efficiency Agreement for Industries	Reporting obligations and the monitoring system were planned at the same time as the policies. Measures have been taken to reduce errors and uncertainty in the data submitted by the participants, by providing them with guidelines, training and support. The participants have a possibility to report on their overall satisfaction to the agreement scheme in their annual reporting. Additional feedback is collected in various events organized around the agreement scheme. Feedback was also the topic of the formal third-party evaluation done in 2005.
[FI] Voluntary energy audits for municipalities	Training and qualification schemes for energy auditors, to ensure quality of energy audits and reporting. A web platform enables an easy monitoring of the energy audits and actions recommended. Before this platform was available, participating municipalities reported annually the status of the actions recommended in the energy audits. The web platform has made possible an update of the status of the actions at any time (easier to manage than a single update per year). Motiva makes plausibility check of the data reported. And a sample of actions is verified each year. "We used our knowledge gathered through our long history of M&V (1994->) of voluntary energy audits when we planned M&V for the mandatory audits."
[FR] Voluntary agreement for freight companies	Since 2015, the companies willing to get the (voluntary) scheme label must be audited by an independent auditor. During the on-site audit, a particular attention is paid to the procedures for data collection, the calculation of the fleet performance (validation of the calculation methodology used for the reference period) and the documentary proficiency/procedures (among other aspects).
[FR] "Future Investments" programme	The mid-term evaluation made possible to improve the procedures used to review and contract the proposals applying for public aids, and to improve the tools to facilitate data collection from project holders in order to provide more reliable data for the monitoring and evaluation of the projects.

Case study	Good practices about data collection and monitoring
[DE] Energy Efficiency Networks Initiative	Participating in the monitoring process is one of the requirements for the networks to be registered. "The two-step evaluation system of monitoring at the administrating agency and an independent evaluation has proven successful." (the scheme administrators monitors the targets defined by each network, and independent institute evaluates the target achievement) "The yearly conducted telephone survey has shown that the monitoring and the connected effort with supplying the data is becoming less of a barrier than it used to be in the beginning of the scheme."
[DE] Energy Efficiency Fund	The evaluation of an umbrella programme like the Fund (that includes of a range of measures) is challenging to ensure that the evaluation of each measure follows the same methodology for consistency purposes. This is particularly true when different evaluators are in charge of evaluating different measures. For this purpose, a very detailed methodology report was elaborated first detailing which data is necessary for an aggregated evaluation. Indicators are clearly defined and numbered. Methodologies for aggregating emissions and energy savings, are also described in detail.
[IE] Better Energy Homes	Several ex-post studies were made to progressively get a better picture of the impacts of the scheme. These studies provided a feedback used to improve data collection done for monitoring the scheme.
[IT] White Certificates Scheme	Over time, stricter rules have been set about the responsibilities of actors applying for certificates, making both ESCos and end-users responsible in case of problems leading to penalties decided by GSE. This is important as small ESCos may run big projects for large end-users, not always having financial capacity to cover these risks.
[LT] Renovation programmes with EU funding	By law, each energy efficiency project receiving a public aid should report data, and data requirements are defined in national monitoring rules.
[NL] Multi-year agreements in the industry	"Better incentives are needed to ensure that companies provide reliable data. A good set-up of an energy efficiency instrumentation requires that one thinks thoroughly through how it will be evaluated."
[NL] Subsidy scheme for housing corporations	"In a subsidy scheme you can make rules that parties that receive subsidies are obliged to make information available and accessible for evaluation. So you can make sure that the design of the policy is evaluative: goals are clearly formulated, data to measure the effects should be available, etc."
[Nordic Countries] Nordsyn	The cost of market surveillance would have been higher if the synergy between countries and their common surveillance was not utilized. This is one of the main conclusion of the evaluation: the countries need to act together to use the best of the market data and to yield the largest savings through greater compliance. "cooperation among MSAs [market surveillance authorities] and policy experts of different countries creates knowledge and possibility for all to advance"
[UK] Supplier Obligations	"You need to think about evaluation when you start the scheme, embed it in the design and make sure data is collected from the start. This can only happen if there is a legal framework for suppliers to actually provide the data." NEED (National Energy Efficiency Data-framework) is the largest source of data available in the UK for analysis of consumption and the impacts of installing energy efficiency actions. It combines data from existing sources (administrative and commercial) to provide insights into how energy is used and what the impact of energy efficiency actions are on gas and electricity consumption, for different types of properties and households. The data framework matches gas and electricity consumption data, collected for BEIS sub-national energy consumption statistics, with information on energy efficiency actions installed in homes, from the Homes Energy Efficiency Database (HEED), Green Deal, the Energy Company Obligation (ECO) and the Feed-in Tariff scheme. It also includes data about property attributes and household characteristics, obtained from a range of sources.
[US] Auctions for capacity markets	"Flexibility is needed for M&V as projects are diverse and keep changing with innovation, which makes M&V challenging to an extent. The right balance has to be struck between strict M&V rules to ensure reliability and the need for flexibility to accommodate for the diversity of projects." The requirements for demonstrating successful delivery of calculated capacity reductions are substantial. One utility reported that up to 30% of the revenue received in the capacity auctions is taken up in the administrative costs of participating in the auctions and demonstrating compliance. However, it should be noted that revenues from the capacity auctions can be used as complementary

Case study	Good practices about data collection and monitoring
	funding to strengthen energy efficiency activities or to reduce the levy on customer energy bills.
[US] Weatherization Assistance Program	"Another lesson learnt is that one should not overburden with requests the people whose activities are evaluated. Due to ARRA, WAP budget was very largely increased. This made that state agencies and Sub-Grantees were already very busy with growing activities when they were contacted to provide data." "If procedures are put in place in advance such as ... #1 Development of a good quality tracking database; #2 Ensuring the clients sign utility waivers that have been accepted by the state's utilities (taking into account state's specificities); #3 Making monitoring data accessible for analysis ... state level studies can be done quickly and efficiently to give DOE, the state grantees, and the agencies important information about program performance at a relatively low cost. A number of states ... Iowa, Illinois, and Wisconsin ... have conducted annual program evaluations that are quite informative and lead to program improvements."

5 | Evaluation practices: what methods and what for

The main conclusion from this part is:

MESSAGE 06: Regular review and in-depth ex-post evaluations are complementary.

The analyses done in the case studies led to distinguish two main practices of ex-post evaluation:

1. Regular reviews, usually done annually and based on the verification and compilation of data from on-going monitoring and/or annual reports prepared by participants (voluntary agreements) or obligated parties (energy efficiency obligation schemes).
2. In-depth ex-post evaluations, usually covering a multi-year period, and including further data collection (e.g., surveys, interviews) and analysis.

This is illustrated in the information collected below. Sections 0 and 5.2 details the methods used for annual reviews and multi-year evaluations respectively. Section 5.3 then gives an overview summarizing for which policy measures regular reviews and multi-year evaluations have been used, showing that most often both are used. Which supported the conclusion about their complementarity. Finally, section 5.4 provides practical examples of this complementarity.

As highlighted in part 4 of this report, the review of monitoring and evaluation practices shows that monitoring is most often a pre-requisite for evaluation, providing the data used as starting point for further analyses done in ex-post evaluations. These further analyses are usually done for two main reasons:

1. Providing an external look at the policy measure, and more specifically about how it works and could be improved.
2. Investigating the causality between the policy measure and its results, or in other words assessing the net impacts of the policy measure.

In most of the cases, the monitoring systems provide data that are used to calculate gross energy savings, i.e. energy savings from all actions for which data were collected, not including adjustment effects such as free-rider or spill-over effects. In a few cases, special factors are included in the calculations to take into account possible correction or adjustment factors. But in these cases, these factors are defined beforehand, either based on literature or on previous evaluations (but not based on effects observed for the period monitored). In some cases, additionality criteria are used to define the baseline or what actions are eligible. But like for the use of special factors, these criteria are defined beforehand, mostly based on assumptions about what the effects of the policy measure can be, or to avoid double counting with other policy measures.

That is why ex-post evaluation can bring an added value in the understanding, and when possible quantifying, of the net impacts of the policy measure. This is further illustrated in part 7 of this report.

5.1 Methods used for regular reviews

NOTE: the analysis here is focused on the calculation methods used to assess the energy savings reported for official purposes (e.g., annual report of the policy measure, reporting to the European Commission).

Table 12. Methods used for regular reviews (official reporting).

Case study	Evaluation method	Comment / quotes
[AT] Environmental Support scheme	Method 5 (scaled savings)	Ex-ante + systematic verification of the calculations when processing the applications. Use of scaled savings (engineering calculations), as projects can be very specific (particularly for the industry sector) No correction factors used (calculations take into account normalised operating conditions when needed) Projects are additional (cf. additionality requirements) but savings are "gross" savings, as baseline = "actual before".
[AT] City Energy Efficiency Programmes of Vienna	Method 3 (deemed savings) or Method 5 (scaled savings)	Use of the calculation methods established by the national monitoring body in the frame of the Energy Services Directive (2006/32/EC). SEP savings are additional compared to minimum energy performance requirements enforced by EU and Austrian regulations. This way, double counting with Federal measures is avoided.
[BE] Primes Energie	Method 5 (scaled savings)	Bottom-up calculation methodology based on the EC recommendations. Engineering calculations combining data monitored and specific to the actions (energy performance improvement) and reference values (baseline situation). Regular update of the reference values (based on regional statistics, databases of energy audits and EPC). Use of normalised weather conditions and behaviours, no other adjustment factor applied. Performance criteria on actions to ensure performance additionality, but savings are gross savings as the baseline is the stock average.
[CR] Energy renovation of public sector buildings	Method 5 (scaled savings)	Ex-ante assessment (engineering simulation) registered online (SMiV platform). Normalization of weather conditions, occupancy rates and operating hours. Gross savings (no causality assessment).
[CR] Individual heat metering in multifamily buildings	Method 3 (deemed savings)	Default value set in the legislation (10% savings) Normalization of weather conditions. No assessment of rebound effect, but it is noted that most buildings were over-heated before the installation of heat allocators (cf. collective heating without individual metering)
[DK] EEO scheme	Method 3 (deemed savings), mostly for residential sector or Method 5 (scaled savings), mostly for industry and services	Annual reporting by the obligated parties, using either scaled savings (mostly for actions in industry) or deemed savings (mostly for actions in the residential sector) depending on the type of actions and sector. Monitoring and verification done by the Danish Energy Agency. Deemed savings are normalized (e.g., weather conditions, heating behaviours). Scaled savings are corrected for changes in operation hours, production volumes, etc. Conversion factors (for substitution between energy sources), reduction factors (based on additionality assessments done in previous ex-post evaluations), and prioritisation factors (to favour some action types, e.g. actions with longer lifetime). Reported savings are thus additional savings, based on how additionality is defined.
[FI] Energy Efficiency Agreement for Industries	Method 5 (scaled savings)	Detailed engineering estimates often made by external consultants for the participants. Double counting with other policy measures is tracked.
[FI] Voluntary energy audits for municipalities	Method 5 (scaled savings)	Detailed engineering calculations made by the energy auditors. Then implementation of the recommended actions is monitored through the reporting of municipalities, as part of the voluntary agreement.

Case study	Evaluation method	Comment / quotes
[FR] Voluntary agreement for freight companies	Method 8 (Monitoring of energy consumption indicators)	Calculation based on energy consumption indicator per ton.km or per passenger.km (including normalisation for distances travelled and weight or passengers carried). Data about energy consumption, distances travelled and passengers or weight carried are reported by the participant companies (either based on measurements or estimations). Data collection procedures and calculations are verified by an external auditors for the companies willing to get the scheme label.
[FR] "Future Investments" programme	Method 5 (scaled savings)	In principle, energy savings are to be assessed by multiplying the potential energy savings due to the supported technology (compared to a reference scenario) by the expected number of technologies assessed for the "market development" scenario designed to size the conditions for the aid reimbursement. This ex-ante estimate is prepared by the project holder and verified by experts of ADEME.
[DE] Energy Efficiency Networks Initiative	Method 3 (deemed savings)	Monitoring is based on the targets defined by each network.
[DE] Energy Efficiency Fund	Method 3 (deemed savings) or Method 5 (scaled savings)	Gross final first year energy savings are calculated before the implementation of the action, based on the information provided by the beneficiaries (Method 3 / deemed savings) or the energy auditors (Method 5 / scaled savings).
[IE] Better Energy Homes	Method 4 (deemed savings)	Reported energy savings are based on the regular monitoring (using simplified engineering calculations). They do not include energy savings from boiler replacements (reported for the building regulations) and are gross energy savings, taking into account rebound effect (but not possible free-rider effects) Use of normalised weather conditions. Rebound effect (conservative values per type of dwelling, based on the comparison between modelled and metered energy consumption). Default value for free-rider effect taken into account in the cost-benefit analysis, but not in the reported energy savings.
[IT] White Certificates Scheme	Method 4 (mix of deemed savings and metered savings), methods 1 or 2 (metered savings)	Since 2017, two methods can be used to evaluate the savings from a project: For standard projects, savings are calculated based both on the installed units and the measurements done on a statistically representative sample. For monitoring plans projects, savings are measured on the basis on one or more meters. The energy consumption baseline and the additionality are determined on a project-to-project basis, taking also into account normalisation (e.g. for manufactured volumes, plants usage, weather, etc.) and market average (to set the baseline). Double counting (verifying certificates are not issued twice for the same action).
[LT] Renovation programmes with EU funding	Method 5 (scaled savings)	Energy consumption before and after renovation is calculated using the methodology defined for the energy certification for buildings, i.e. assuming standardized heating behaviours and weather conditions. No adjustment (rebound effect, free-rider effect, etc.) is applied. So the results are gross energy savings. The Calculation Rules are defined in the "Rules for calculating national energy savings", approved by Order No 1-320 of the Minister for Energy (05 December 2016)

Case study	Evaluation method	Comment / quotes
[NL] Multi-year agreements in the industry	Methods 3 and 4 (deemed savings)	The baseline is equivalent to the energy consumption before the action is implemented ("actual before"). For MJA3, the reference year is 2005. Only new energy-saving actions newly implemented or intensified from 2005 onwards are taken into account. The energy savings in the annual reports are gross energy savings: they do not include any causality assessment (i.e. possible free-rider or spill-over effects are not taken into account)
[NL] Subsidy scheme for housing corporations	Method 3 (deemed savings)	The monitoring of the scheme used a default value of energy savings per label step.
[NL] Fiscal incentives for cars	Method 3 (deemed savings)	Monitoring of the trend in the specific emissions of new cars (in gCO_2/km), based on the energy labelling for cars (based on normalised tests of cars)
[Nordic Countries] Nordsyn	Method 1 (metered savings)	Tests of samples of appliances in laboratories, done by each national Market Surveillance Authority and gathered by the Nordsyn cooperation.
[UK] Supplier Obligations	Method 3 (deemed savings)	pre-defined carbon/ energy savings ratios according to standardised types of actions taking into account in-use factors which account for rebound effects and performance gaps
[UK] Warm Front	Method 5 (scaled savings)	Annual results reported by the scheme administrators based on simplified engineering calculations done for each participating dwelling, using the SAP (Standard Assessment Procedure) methodology (building energy rating).
[US] Auctions for capacity markets	Method 1 (metered savings), method 4 (deemed savings) or method 5 (scaled savings)	Deemed (method 4) and scaled (method 5) savings can be used if complemented with metered data. The main indicators monitored are the capacity cleared and the verified capacity performance. Energy savings of the projects included in the bids are not evaluated by the capacity market scheme. However, as most of EE resources in the forward capacity market are part of a regulated utility EE obligation, their energy savings are evaluated under the regulatory framework of utility EE obligation instead.
[US] Weatherization Assistance Program	Method 5 (scaled savings)	Monitoring is based on the data reported by the grantees (states) from the data collected by the sub-grantees (local implementing agencies). These data are based on computerized audits done for each home weatherized.

5.2 Methods used for multi-year evaluations

NOTE: the analysis here is focused on the calculation methods used to assess energy savings, and on other evaluation methods used in official evaluation reports (i.e. done or commissioned by the public bodies in charge of the policy measure). It does not include the evaluation methods used in other evaluation studies done on the policy measure.

Table 13. Methods used for multi-year evaluations.

Case study	Evaluation method	Comment / quotes
[AT] Environmental Support scheme	Gathering and verifying data from annual reporting + complementary interviews to analyse specific issues.	Plausibility check in order to make the effects consistent with the overall energy and greenhouse gas emissions balances of Austria.
[AT] City Energy Efficiency Programmes of Vienna	Review of the bottom-up results reported for each scheme of the programme. Top-down approach to monitor the trends in the total energy consumption.	plausibility check of the bottom-up results by comparing with the top-down analysis

Case study	Evaluation method	Comment / quotes
[BE] Primes Energie	Policy portfolio: summing up the impacts of the different policy measures, taking into account only actions that received a financial incentive to avoid double counting. This overall result was then compared to the trends observed in final energy consumption of Wallonia. Umbrella policy framework: evaluation was based on the databases used to monitor the measures included in the umbrella framework, econometric analyses and qualitative surveys of both, participants (households) and building companies	Policy portfolio: evaluation focused on energy savings (+ ex-ante evaluations also done to estimate future impacts according to different assumptions about trends in the number of actions implemented) Umbrella policy framework: evaluation focused on socio-economic impacts (including assessment of free-rider effects for some of the policy measures)
[CR] Energy renovation of public sector buildings	Ex-post evaluation mostly based on data collected by the on-going monitoring (SMiV), that are further analysed	Evaluation performed as new plans are being developed, with analysis focusing on cost effectiveness. "Detailed analysis and evaluation of the programme, alongside with identification of key obstacles, and lessons learned is performed within development of new programmes and plans."
[CR] Individual heat metering in multifamily buildings	Unitary energy savings established on the basis of billing analysis (method 2, metered data + weather normalization)	First evaluation study (2016): sample of 56 buildings (connected to district heating) in 8 cities totalling 3,842 dwellings. Focus on applicability of individual heat cost allocators to Croatian context and identifying common problems. Second evaluation study (2017): larger sample of 276 buildings (22,475 dwellings), from the same 8 cities. Focus on economic feasibility and cost-efficiency.
[DK] EEO scheme	Verifications of data reported + surveys (about additionality) (including comparisons between participants and non-participants) + billing analysis (method 2) on samples (2012) + statistical methods for sectoral analysis (method 9, econometric modelling) (2015)	In 2012, the billing analysis was done as an experimentation (samples non-representative) In 2015, the sectoral analysis was conclusive only for the industry sector (results not statistically significant for the residential sector). Debates about the results from surveys (for additionality).
[FI] Energy Efficiency Agreement for Industries	Review of the information available on the scheme, and interviews with participants and other stakeholders.	Third-party evaluation done once (in 2004-2005)
[FI] Voluntary energy audits for municipalities	Basis is the annual monitoring & evaluation. No particular further evaluation	Particular attention paid to avoid double counting with other policy measures.
[FR] Voluntary agreement for freight companies	Review of the data reported by the participant companies, complemented by an online survey of participants	Ex-post evaluation done after 8 years of implementation (including a process evaluation including interviews with stakeholders and regional case studies)
[FR] "Future Investments" programme	Econometric evaluation of the effects of the aid scheme on beneficiaries (R&D expenses, patents delivery, sales, staff employment…): use of the Difference-in-Differences method (similar to method 2); Quantitative-qualitative evaluation of the results and impacts of supported projects and of the programme as a whole	Mid-term external ex-post evaluation The DiD method was used to look at direct impacts on participants' achievements of environmental objectives, investments in R&D and in environmental fields + indirect impacts on participants' business development (production, employment, productivity, etc.) and leverage effect

Case study	Evaluation method	Comment / quotes
[DE] Energy Efficiency Networks Initiative	Achieved energy savings are evaluated after network operation time, based on data reported by the participants (detailed survey). Data reported by participants can include metered energy savings (methods 1 or 2), deemed savings (methods 3 or 4) or scaled savings (method 5). Data are checked by the evaluators, including detailed review of documentation for 10% of the companies.	The participant companies are surveyed for their energy savings using an extensive questionnaire including description of energy efficiency actions, status before implementation, savings calculation and baseline calculation, energy source with reduced consumption, newly implemented energy source in case of a change in energy sources.
[DE] Energy Efficiency Fund	The EEF is re-evaluated and updated each year until its end in 2020.	The individual policy measures as well as the fund as a whole are evaluated by independent entities regularly using both qualitative and quantitative approaches depending on the measure. Free-rider effects determined based on ex-post surveys. Double counting is tracked (interaction effects between the different sub-measures of the Fund)
[IE] Better Energy Homes	difference-in-difference method (quasi-experimental approach), comparing pre-(2008) and post-intervention (2010) heating consumption for participants and control group (method 2, billing analysis)	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?
[IT] White Certificates Scheme	Review of the annual reports prepared by the obligated parties, and of the results for the verifications and controls done by GSE.	Review used to monitor target achievement, and to see if the scheme needs to be adapted/revised.
[LT] Renovation programmes with EU funding	For some selected residential buildings, ex-post monitoring (or energy efficiency audits) on implemented actions and energy savings is performed.	
[NL] Multi-year agreements in the industry	Further data collection (literature review, interviews and survey). A decomposition analysis (method 9) was done to show, among others, the changes in energy consumption due to factors other than the energy savings project, such as changes in factory design and product specifications changes in the production volume, changes in the ambient temperature, etc.	Ex-post evaluation done in 2012-2013 about the period 2008-2012.
[NL] Subsidy scheme for housing corporations	The evaluation done by the Amsterdam audit office used a stock modelling (method 6). The further investigation done by the Technical University of Delft used a comparison between energy labels and energy bills (metered savings, method 2).	The evaluation by the audit office was mainly based on data from a very large sample of dwellings that can be taken as a representation of the dwelling stock (this sample was not made from dwellings that were renovated). Whereas the further investigation by the Technical University of Delft used data from the dwellings that were renovated.

Case study	Evaluation method	Comment / quotes
[NL] Fiscal incentives for cars	The evaluation done by PBL included a stock modelling (method 6) to assess the effects of the policy on the sales (and consequently stock) of cars, as well as an econometric analysis to investigate to what extent the tax measure stimulated the sale of clean and efficient cars (causality), and leakage effects (indirect impacts on car sales in other countries). In parallel, test procedures were used to analyse fuel consumption and CO_2 emissions (difference in test lab results vs. results in the field (method 1))	The trend in the CO_2 emissions of new cars in the EU countries was used as a benchmark to define the baseline for the impact evaluation.
[Nordic Countries] Nordsyn	Method 4 (deemed savings, with a mix of metered and estimated data) The evaluation gathered all data available from the market surveillance authorities (MSA) partners of Nordsyn and did complementary assumptions to estimate the potential of "lost" energy savings in the whole countries involved in Nordsyn.	The approach was to evaluate the energy impacts of market surveillance by estimating the over-consumption of non-compliant products based on the results from the tests on samples of appliances and sales data. Key points are the estimation of the non-compliance rates per type of appliance, taking into account the sampling method used by each MSA. Free-rider effects do not apply to Nordsyn, as it is about the implementation of a regulation. Spill-over and rebound effects were not taken into account in the evaluation.
[UK] Supplier Obligations	The National Energy Efficiency Data-Framework (NEED) was set up by BEIS to provide a better understanding of energy use and energy efficiency in domestic and non-domestic buildings in Great Britain. NEED enables analysis similar to billing analysis (method 2).	The UK government has commissioned multiple studies in order to derive and improve the deemed savings estimates. Those studies have been carried out by independent consultants, academics and the Energy Saving Trust.
[UK] Warm Front	The study done about health impacts investigated the changes in energy consumption through measured data in samples of participants and in a comparison group (method 2, billing analysis). The study done by BRE was equivalent to a stock modelling (method 6), using national statistics to model the dwelling stock and simulating scenarios to assess the impacts of the scheme.	Several studies were done to assess the differences between the theoretical energy savings monitored and the actual energy savings. The BRE study applied a reduction factor, made of a 40% "comfort factor" (assumed direct rebound effect) and, in the case of the insulation actions, of a 41% "underperformance factor". These assumptions were based on previous studies.
[US] Weatherization Assistance Program	Energy bills analysis (method 2, metered energy savings): pre/post treatment/comparison statistical analysis (statistical method of Difference in differences), using weather normalized utility billing (on samples then extrapolating results using weighting factors by housing unit type) In addition, an Indoor Environmental Quality (IEQ) study about works completed during the ARRA period concluded that there was no rebound effect relevant to home heating.	"The evaluations done in 2008 and 2011 represented an about $20 million effort, which made possible to have a dedicated data collection. One of the objectives was to verify the reliability of the monitoring. That's why it could not be based on monitoring data." Gross energy savings were evaluated by comparing energy consumption before and after the WAP intervention ("actual before" baseline). Net energy savings were evaluated by comparing changes in energy consumption between the samples of participants and control groups ("control group" baseline).

5.3 Overview of the organisation of evaluation practices

Table 14. Overview of the use of regular reviews and multi-year evaluations in the EPATEE case studies.

Case study	Annual	comment	Multi-year	comment
[AT] Environmental Support scheme	X	Calculations checked for every application before its approval + annual review of the results	X	External evaluation every 3 years. Further verifications including plausibility check to make the effects consistent with the national energy and greenhouse gas emissions balances + complementary interviews to investigate specific issues.
[AT] City Energy Efficiency Programmes of Vienna	X	Monitoring and ex-post verification is done separately for each scheme of the umbrella programme. The level of ex-post verification is highest for subsidy schemes.	X	External ex-post evaluation every three years, combining a review of the bottom-up results and top-down analysis
[BE] Primes Energie	X	Monitoring done by the service in charge of the scheme, with regular updates of the reference values (baseline)	X	Internal evaluation done by another service (frequency based on reporting cycles)
[CR] Energy renovation of public sector buildings	X	Continuous monitoring based on scaled savings	X	When developing a new programme or plan (but mostly based on the same data as the monitoring)
[CR] Individual heat metering in multifamily buildings	X	Continuous monitoring based on deemed savings	X	Two ex-post studies (2016 and 2017) to investigate actual energy savings and cost-effectiveness
[DK] EEO scheme	X	Annual reporting, with verifications on samples	X	External ex-post evaluation at the end of each period
[FI] Energy Efficiency Agreement for Industries	X	On-going monitoring used to produce annual reports	X	Third-party evaluation done once (in 2005)
[FI] Voluntary energy audits for municipalities	X	On-going monitoring used to produce annual reports		No further evaluation study, but compilation of data (for example to report to the European Commission)
[FR] Voluntary agreement for freight companies	X	Monitoring based on data reported annually by the participant companies (+ external audits of the companies willing to get the scheme label)	X	Done once, after 8 years of implementation of the scheme About impacts, mostly based on the review of data submitted by participant companies for the monitoring (additional data collected about costs and for qualitative aspects / process evaluation)
[FR] "Future Investments" programme	X	ex-ante evaluation of each project (technical, economical, financial & regulatory) carried out by experts from ADEME to size the financial support.	X	ex-post programme evaluation requested by the CGI (the French governmental organization supervising the programme) and the European Commission to comply with the EU obligation for large state aid programme
[DE] Energy Efficiency	X	monitoring of the targets defined	X	Evaluation of target achievement at the

Case study	Annual	comment	Multi-year	comment
Networks Initiative		by each network (ex-ante data)		end of commitment period of each network. Overall evaluation of the impacts of the scheme will only be possible in 2020.
[DE] Energy Efficiency Fund	X	Monitoring and review based on the data submitted by the applicants before the implementation of the projects.	X	The individual policy measures as well as the fund as a whole are evaluated by independent entities regularly using both qualitative and quantitative approaches depending on the measure.
[IE] Better Energy Homes	X	Regular reporting	X	Complementary impact evaluation (in 2009), to investigate main source of uncertainties, and to get more robust results. + participants' survey to assess satisfaction and how the scheme could be improved
[IT] White Certificates Scheme	X	Systematic review of the documentation of the projects, and annual review of obligated parties' report		
[LT] Renovation programmes with EU funding	X	Ex-ante energy audit done for each project (+ building energy certificates before/after), and data centralised by the Housing Energy Efficiency Agency		Ex-post monitoring (or energy efficiency audits) on samples of implemented actions (but this is not included in an official and formal evaluation of the programme)
[NL] Multi-year agreements in the industry	X	Annual review based on data reported by the companies that signed the agreement	X	Ex-post evaluation done in 2012 about the 2008-2012 period
[NL] Subsidy scheme for housing corporations	X	Monitoring of the scheme by the municipality of Amsterdam (mostly based on the number of label steps reported by the housing corporations)	X	Ex-post evaluation done by the Amsterdam audit office in 2014 (own initiative), with an approach similar to a stock modelling.
[NL] Fiscal incentives for cars	X	Monitoring of the sales or cars and corresponding energy labels	X	Ex-post evaluation done by PBL in 2014 (own initiative), with an economic modelling to assess changes in the stock of cars, complemented by test procedures for analysing CO_2 emission (difference in test lab results vs. field measurements)
[Nordic Countries] Nordsyn	X	Laboratory tests on samples of appliances	X	Evaluation study done in 2014-2015. Gathering data from all market surveillance authorities and doing complementary assumptions to extrapolate the results to the whole markets of the countries partners of Nordsyn.
[UK] Supplier Obligations	X	Monitoring is done on a monthly basis	X	"The Supplier Obligation is continuous and there is typically not enough time to carry out an evaluation at the end and apply the learnings in the next phase. We therefore conduct post implementation reviews which happen mid-term during an ongoing phase of the Supplier Obligation."
[UK] Warm Front	X	Annual reporting by the scheme	X	External studies commissioned by the

Case study	Annual	comment	Multi-year	comment
		administrators based on the monitoring of the scheme		ministry to look at health impacts (and intermediate impacts such as actual changes in energy consumption) (2001-2005) and to assess energy and CO_2 savings with a more sophisticated modelling than done in the annual reports (2008).
[US] Auctions for capacity markets	X	Evaluation of the offers received for the annual auctions, and ex-post verification.		
[US] Weatherization Assistance Program	X	Annual review of the results and performance indicators of the scheme	X	Ex-post evaluations done depending on the needs, since 1989.

5.4 Practical examples about complementarity between annual reviews and multi-year evaluations

Table 15. Feedbacks about complementarity between annual reviews and multi-year evaluations.

Case study	Examples and quotes about complementarity between annual reviews and multi-year evaluations
[AT] Environmental Support scheme	The annual reviews provide a regular feedback loop, while the external evaluations done every three years provide an external look and enable to investigate more in details trends, changes and specific issues identified with the annual reviews. The external evaluators also use the data from the monitoring to perform a plausibility check with the overall energy and greenhouse gas emissions balances of Austria.
[AT] City Energy Efficiency Programmes of Vienna	The evaluation is used to review the data collected about each scheme (bottom-up results). This review is complemented by top-down analysis of the trends in total energy consumption, making possible a plausibility check of the bottom-up results. The external evaluation also provides a view on the effectiveness of each measure and the achievement of the overall target.
[BE] Primes Energie	The database used to monitor the financial incentives and centralise the data provides the input data for the energy savings calculations). The ex-post evaluation of energy savings is therefore mainly based on the data collected through the monitoring system. The ex-post evaluation of the umbrella policy framework has a broader scope and provides a more global view about the impacts and interactions of the different schemes.
[CR] Individual heat metering in multifamily buildings	The ex-post studies were made to complement the monitoring of the scheme, by assessing the actual energy savings and cost-effectiveness of the actions. They were commissioned to investigate under which conditions individual heat allocators are cost-effective in view of updating the regulation.
[DK] EEO scheme	"The ex-post evaluations are used to complement the monitoring of the scheme when preparing a revision of the agreement for the scheme." "It is important to distinguish M&V and evaluation. M&V provides data and feedback as a regular basis for managing the scheme. Evaluation provides an independent and in-depth analysis of the scheme and its impacts, in order to draw recommendations."
[FI] Energy Efficiency Agreement for Industries	"The system works well and it appears that there is not much need for improvement anymore. There is experience already since 2000 and improvements have been made constantly in small steps in the spirit of continuous learning."

Case study	Examples and quotes about complementarity between annual reviews and multi-year evaluations
[FR] Voluntary agreement for freight companies	Data from the monitoring were the basis for the ex-post evaluations that did further analysis of these data to provide an overall estimate of the impact of the scheme. Moreover, the ex-post evaluation also brought complementary data about costs, as well as about qualitative aspects.
[FR] "Future Investments" programme	Monitoring is based on the evaluation of each project. Whereas the ex-post evaluation is meant to assess the overall impact of the scheme.
[DE] Energy Efficiency Networks Initiative	The data monitored about the targets defined by each network provides the point of comparison for the evaluation to assess the effectiveness of the scheme (target achievement).
[DE] Energy Efficiency Fund	The regular monitoring provides the basis for the external evaluations that include further analysis and surveys (for example to assess free-rider effects and net impacts).
[IE] Better Energy Homes	"While engineering estimates are useful to monitor the results on an on-going basis, I strongly recommend to go beyond engineering estimates. One may have fear to do an ex-post impact evaluation, because it may show smaller results than based on the engineering estimates. However this increases the robustness of the results and therefore the confidence funders can have in them. This can be combined with a Cost-Benefit Analysis to show that despite the energy savings being possibly smaller, the overall result for society remains a net benefit, when taking into account all the impacts, beyond energy savings alone. Qualitative analysis is also essential, for example to know how the participants feel about the improvements of their dwelling. This should be combined with the quantitative impact analysis, in order to understand how to promote the scheme."
[NL] Multi-year agreements in the industry	The ex-post evaluation provides further data collection and analysis about the impacts and costs of the scheme, as well as about how it can be improved.
[NL] Subsidy scheme for housing corporations	The ex-post evaluation was done on the own initiative of the audit office, to assess the actual energy savings from the scheme (whereas the monitoring has used deemed savings based on a default value per label step).
[NL] Fiscal incentives for cars	The data monitored provided the key input data for the economic modelling to assess the impacts of the tax measure on the sales of cars, and thereby on the stock of cars.
[Nordic Countries] Nordsyn	Market surveillance activities are a form of ex-post verifications, particularly the laboratory tests that provide part of the key data. It was then complemented by a specific study (the Effect project) to assess the impacts and perform a cost-benefit analysis of these activities. Each market surveillance authority (MSA) uses different sampling methods to select the products to test. A key issue is that the choice of the most appropriate sampling method is not the same depending on the objectives. Market surveillance's priority is to detect and remove non-compliant products, not to have a representative picture of the market. Whereas an impact evaluation would ideally be based on samples as representative as possible, for example using random or semi-random sampling (to enable robust statistical analysis). In practice, MSAs are indeed more often using other sampling methods with a risk-based approach.
[UK] Supplier Obligations	Data from the monitoring provides the starting point of the different evaluations and studies done. Reciprocally, the further analyses done or commissioned by the ministry are used to revise the deemed savings used to monitor the results of the scheme.
[UK] Warm Front	"Choosing the most appropriate approach should consider the constraints (budget and time) and the evaluation objectives. What would be the most useful for the evaluation customers/recipients: knowing the impacts? or understanding the impacts?"

Case study	Examples and quotes about complementarity between annual reviews and multi-year evaluations
[US] Weatherization Assistance Program	So far the data collected for monitoring the scheme have mostly been used to select the dwellings to be included in the samples analysed for the ex-post evaluation. The DOE WAP managers examines the potential of working with the Grantees to increase the usefulness of their databases for future evaluations. Evaluation results are used to improve the monitoring (update of the software for the audits, update of the assumptions used when assessing the overall annual results of the scheme, etc.). "WAP continues to evolve every year. So one should be careful in using past evaluation results to inform current management of the scheme. When applying evaluation results, changes already made should be taken into account. For example to see if there is a need to re-evaluate results used for monitoring."

6 | Evaluation practices: feedbacks about the main categories of methods

The main conclusions from this part are:

MESSAGE 07: The choice of evaluation methods depends on evaluation objectives but also on practical constraints.

MESSAGE 08: There is no silver bullet. All methods include uncertainties. Comparing different methods helps assessing the robustness of the results, and getting a sound basis for decision-making.

This is illustrated in the information collected below. Sections 6.1 and 6.2 gather the feedback collected in the case studies about the use of engineering methods and statistical methods respectively. This shows that there is no "perfect" evaluation method. All methods have their pros and cons, added value and limitations. Section 6.3 then provides examples where several methods have been compared, used to test the plausibility/reliability of the results or to get complementary analysis.

6.1 Feedback about engineering methods

Table 16. Feedback about engineering methods.

Case study	Feedbacks about engineering methods
[AT] Environmental Support scheme	Use of scaled savings (engineering calculations), as projects can be very specific (particularly for the industry sector).
[AT] City Energy Efficiency Programmes of Vienna	The uncertainties of the bottom-up methods (mostly deemed savings and engineering estimates) are connected with the underlying ex-ante approach and the need to make certain assumptions on consumer behaviour (incl. room temperature), climate conditions and the technically optimal installation of technologies. In addition not all actions implemented in the framework of the SEP can be evaluated bottom-up.
[BE] Primes Energie	"It is also an issue to track the changes that may affect the definition of the baseline situations, and to update baselines consistently." "We also encountered some difficulties to get a common understanding of the evaluation of energy savings between experts using different methodologies, due to distinct evaluation objectives. This was the case about the way to estimate the baseline energy consumption. For the evaluation of energy savings from boiler replacements, we used statistics based on metered energy consumption. Whereas the experts in charge of the scheme for Energy Performance Certificates (named PEB in Wallonia) are used to conventional energy consumption (i.e. defined according to standardized assumptions on heating behaviours). This may create some confusions and require to explain the results."
[CR] Energy renovation of public sector buildings	Initially, no connection between databases monitoring energy consumption and actions respectively; this was under consideration and could drastically improve calculation's accuracy. In the future, the intention is to use "energy savings as monitored after the renovation".
[DK] EEO scheme	When the target becomes higher and more difficult to achieve, there are higher risks of errors in the calculations reported by the obligated parties (the higher the number of actions to monitor, the more difficult to make systematic verifications, etc.). Consequently, verification processes were reinforced from early 2017 based on the conclusions from the 2015 evaluation and the report from the Court of Auditors.

Case study	Feedbacks about engineering methods
[FI] Energy Efficiency Agreement for Industries	The savings achieved by the energy saving actions are not usually verified by subsequent measurements, since it is most often difficult to make measurements in practice and it generates significant additional costs.
[FI] Voluntary energy audits for municipalities	Quality ensured through training and qualification schemes of energy auditors + plausibility checks of the data reported.
[FR] "Future Investments" programme	Uncertainties inherent to innovative projects: Projects might result in different outputs than initially planned. Difficulties in estimating the development costs for technologies not yet available on the market. Difficulties to define counterfactual scenarios in case of very specific technologies (too limited number of potential customers to find a significant control group). Moreover, energy savings cannot be monitored or verified ex-post on short term (as they will occur after a certain time depending on technology development and time to market).
[DE] Energy Efficiency Fund	Differences in calculation methodologies used by energy auditors (for complex actions) Uncertainties of values from product data sheets (for simple projects, no ex-post verification) Many of the 23 measures of the Energy Efficiency Fund (EEF) are information or education-based. These are not quantifiable and are excluded in the aggregate analysis. Aggregate results, therefore, systematically underestimate the total savings from the EEF.
[IE] Better Energy Homes	"Empirical verifications represent a small budget compared to the whole budget of the scheme. Our experience with the ex-post impact evaluation is that it is really worth it. While engineering estimates are useful to monitor the results on an on-going basis, I strongly recommend to go beyond engineering estimates. One may have fear to do an ex-post impact evaluation, because it may show smaller results than based on the engineering estimates. However this increases the robustness of the results and therefore the confidence funders can have in them."
[IT] White Certificates Scheme	"The complexity of energy efficiency intervention is increasing. In the first period, single component interventions (e.g. boiler, inverter etc.) were predominant, whereas currently sectorial-specific interventions and industrial process reengineering are more frequent. This increases the complexity of energy savings measurement and also the cost of savings."
[LT] Renovation programmes with EU funding	Investigations to compare scaled savings from energy audits and metered savings have shown discrepancies. Also large variations were observed in the sample used to assess metered savings, leading to the conclusion that further investigations on larger samples would be needed.
[NL] Multi-year agreements in the industry	The objective of the bottom-up approach used from 2010 was to be pragmatic, easy to implement, verifiable and testable.
[NL] Subsidy scheme for housing corporations	Data on energy use and energy labels were linked, but there were difficulties because addresses were not written in a uniform way. The investigations done by the Technical University of Delft showed the limitations of using energy labels as a basis to estimate energy savings and the limitations of using average values taken from different dwellings than the ones where the renovation works were done
[NL] Fiscal incentives for cars	Actual fuel savings are lower than theoretical fuel savings (expected from the emission norm) as the norm relates to a test cycle that does not sufficiently reflect real driving conditions.
[UK] Warm Front	The different external studies done about the impacts of Warm Front in terms of energy savings showed the limitations to use simplified engineering calculations (like building energy rating) to assess energy savings from a scheme tackling fuel poverty for which comfort taking is one of the main objectives (to reduce fuel poverty and related health impacts). "There had been a kind of fairy story about the use of computer modelling in the 1980's and 1990's. With the development of computer capacities, there was an increasing use of modelling. However the ability to validate engineering data from modelling does not cope with measured ("real") data. Research of the last decade showed that it is essential to get "real" data to analyse complex systems such as energy use in buildings. Energy consumption cannot

Case study	Feedbacks about engineering methods
	be explained by the description of the technical systems alone. There are interactions between technical systems, and above all between technical systems and occupants. This complexity is difficult to model."
[US] Weatherization Assistance Program	Computerized energy audits conducted do not seem to have been systematically calibrated with actual energy usage for each home-unit (i.e. baseline energy use and amount of energy savings might be overestimated). Some state grantees do require this calibration, whereas other do not. Experience shows that results are much better in states requiring calibration.

6.2 Feedback about statistical methods

Table 17. Feedback about statistical methods.

Case study	Feedbacks about statistical methods
[AT] City Energy Efficiency Programmes of Vienna	The top-down analysis shows the overall achievement of the target. It provides a better understanding of the main factors influencing the development of energy consumption. But it does not show the contribution of the SEP or any other causality. However framework conditions (climate, population growth, financial crisis of 2008/2009) are taken into account in the analysis.
[CR] Individual heat metering in multifamily buildings	A bootstrapping method was used to estimate the density distribution function of energy savings in order to extrapolate the results from the samples monitored to the entire population of buildings.
[DK] EEO scheme	Difficulties to get market data that would enable a quantitative assessment of the impacts on EE markets. Difficulties to get disaggregated-enough data and consistent long-time series to perform econometric analyses per end-use sector: the econometric analyses provide results statistically significant only for the industry sector. "In the early 2000's, there was an evaluation of the previous scheme (obligation of energy advice). It tried to compare a control group of companies with a group of companies that received energy advice due to the scheme. This turned to be very difficult, because it required to collect data at the plant level (while many data are more easily available at the company level) and because long time series were required whereas the way the data are monitored (information systems) changed over time. At the end, the result was that the annual random variations (statistical noise) were too large. So it was not possible with this method to distinguish the possible effects of the scheme compared to the effects of other factors. It would be now even more difficult to use this method, as the scheme has now been in place for many years. So it would be very difficult to define a control group with companies that would not have been involved in the scheme and without selection bias. That's why other methods have been used in the later evaluations." "The introduction of smart meters will provide better access to certain data opening up for alternative evaluation approaches."
[FR] Voluntary agreement for freight companies	Issue of representativeness of the data used to assess the impact: The extrapolation of data collected to estimate the impact of the scheme on CO_2 emission reduction may have biased the results interpretation. Indeed, data of companies involved for only one year were not necessarily representative of their involvement during the entire period of the scheme, and the behaviour of certain companies was potentially biased by the short-term perspective of the involvement (3 years).

Case study	Feedbacks about statistical methods
[FR] "Future Investments" programme	One of the major evaluation challenges is that the evaluation results might be biased by variables other than the programme effects and that affect the observed outcome such as general macroeconomic conditions or firm heterogeneity The number of supported projects for environmental protection and energy identified in 2016 for the mid-term ex-post evaluation was too low to generate a sample statistically robust for an econometric evaluation. The first results have shown that the nature of the sample available for the econometric study underway does not support the interpretation of results as significant causal effects of the programme on the variables of interest (proxies for R&D expenditure, R&D jobs, etc.). The mid-term evaluation confirmed the difficulty in implementing econometric methods in practice, and the difficulty in obtaining robust results. However, it has helped understanding the causal processes leading (or not) to technological and/or commercial successes.
[IE] Better Energy Homes	"We faced some difficulties for the billing study to access billing data. The main issue was that for legal reasons, we had to contact participants to get their agreement to use their billing data for the evaluation (even if these data were anonymised). This took a lot of efforts and led to a smaller sample compared to the initial plan. We therefore strongly recommend to put a condition when setting the scheme to ensure the access to data, and that the legal team checks the validity of this condition. This should be possible for any grant scheme, as participants will very likely agree on this condition to be able to get the grant. This can save a lot of data collection efforts. Another difficulty was related to matching the comparison and participants' groups. This meant to handle a lot of data, which may be technically difficult. And this also requires to be very cautious for respecting the confidentiality rules about data. Billing data were provided by the gas grid operators. Most of them have a small staff, and the requests we send them were on top of their regular job. So efforts were also needed to involve them in the evaluation. Overall, the whole process of data collection and processing took about 2 years."
[NL] Multi-year agreements in the industry	When using the top-down approach (energy efficiency indicators), difficulties were encountered to separate the changes due to the scheme and the changes due to other factors.
[NL] Subsidy scheme for housing corporations	Comparing the evaluation done based on an approach similar to stock modelling (based on data from dwellings that were not the dwellings subject to renovation) and the investigations based on metered consumption of the dwellings subject to renovation showed that the approach based on stock modelling did not reflect the effects of the renovation works.
[NL] Fiscal incentives for cars	Limitation of using an indicator based on a specific energy consumption when it is defined based on laboratory tests (here specific fuel consumption of cars).
[Nordic Countries] Nordsyn	Difficulties due to the different sampling methods used in each country, and the fact that these sampling methods are most often using a risk-based approach, thereby not aiming at representativeness. Assumptions are then needed to extrapolate the results from the samples of appliances tested.
[UK] Supplier Obligations	"NEED [National Energy Efficiency Data-framework] is used as the main tool to assess the actual savings of actions, but there are significant lags between implementation and data availability." For more details about NEED, see: https://www.gov.uk/government/collections/national-energy-efficiency-data-need-framework

Case study	Feedbacks about statistical methods
[UK] Warm Front	"In practice, two constraints make that ideal conditions for a perfectly robust evaluation are very rarely met: money, and often even more importantly time. If we take the example of evaluating impacts on heating consumption. This requires monitoring consumption over at least one heating season before and one heating season after the intervention. Which already means a 2-year experiment. Then a third year is probably needed for the analysis and addressing issues in data collection, etc. But the evaluation customer normal wants results within one year, or less. There is a need to balance between a rigorous statistical analysis with practical time and money constraints. These constraints should be considered when designing the evaluation, defining sample sizes, etc." "With the budget available, it was chosen to monitor temperatures in 1 or 2 rooms per dwelling. This meant the implicit assumption that there was no change in temperature in the other rooms. This is a critical assumption, particularly when evaluating a policy tackling fuel poverty. In-depth interviews indeed confirmed that one of the results of the Warm Front interventions was that households could heat more rooms. So, it would be needed to monitor more rooms per dwelling to capture all the main changes that can affect energy consumption."
[US] Auctions for capacity markets	Market participants need to include a description of the methods used to mitigate and adjust for the potential types of bias resulting from statistical methods related to the accuracy and calibration of the measurement tools, sensor placement bias, and sample selection bias or non-random selection of equipment and/or circuits to monitor.

6.3 Comparing results from different methods or making them complementary

Table 18. Examples of use of different methods.

Case study	Approach / type of use	Comment or quote
[AT] Environmental Support scheme	Plausibility check in order to make the effects consistent with the overall energy and greenhouse gas emissions balances of Austria.	Results are also compared between periods.
[AT] City Energy Efficiency Programmes of Vienna	The evaluation of energy savings uses a bottom-up as well as a top-down approach. The top-down approach has been used to monitor the overall trends in total energy consumption. The bottom-up methods have been used to evaluate the energy savings from each scheme of the programme, only taking into account actions that were triggered by these schemes.	The evaluators highlighted that the bottom-up calculations do not encompass all the effects of the SEP (especially in industry and private services), and that the top-down approach cannot assess the causality between the changes in the overall trend in total energy consumption and the schemes of the SEP. Top-down and bottom-up calculations were combined to check plausibility of energy savings.
[BE] Primes Energie	The ex-post evaluation done in 2014 summed up the impacts of the different policy measures for energy efficiency in buildings, taking into account only actions that received a financial incentive (Primes Energie or other incentive) to avoid double counting. This overall result was compared to the trends observed in final energy consumption of Wallonia.	"Another important point is to combine several data sources for validating the data and key assumptions. This makes possible to have results based on the best data available."

Case study	Approach / type of use	Comment or quote
[CR] Individual heat metering in multifamily buildings	Use of different scenarios to evaluate under which conditions implementation of heat cost allocators would be cost-effective. In addition to these scenarios, the evaluation did an economic sensitivity analysis considering different trends in energy prices and discount rates.	The second evaluation study aimed at determining the cost-effectiveness of heat cost allocators.
[DK] EEO scheme	Different methods used to assess the additionality/net savings: surveys, comparison of a control and participants group (done in 2012), econometric analyses (panel data regression and co-integration) used to assess net effects at sector level (in 2015)	About the 2012 evaluation: "We chose on one evaluation of the obligation scheme to use both a top-down and a bottom-up approach. We used quite a lot of effort to do the top-down assessment. But data were too flawed and the relative impact too small to be discernible. Therefore, the bottom-up approach was preferred by the tenderer for the successive evaluations. However, the roll-out of smart meters creates new opportunities for future evaluations and potentially better access to disaggregated data." "We also routinely look to other evaluations of the same topic to compare results and ask the question whether the evaluations point to the same result or not and why. Especially when sample sizes are small." "In terms of evaluation methods, it is essential to use triangulation, i.e. to compare results obtained with different methods and/or from successive evaluations. This provides a stronger basis for our conclusions"
[FI] Energy Efficiency Agreement for Industries	Plausibility check of the data reported	
[FI] Voluntary energy audits for municipalities	Plausibility check of the data reported	
[FR] Voluntary agreement for freight companies	No comparison made of different methods, but the evaluation included an international benchmarking, with a comparative analysis of seven voluntary schemes for transport companies.	The international benchmarking was meant to put the results from the French programme into a broader perspective and to enrich the recommendations.
[FR] "Future Investments" programme	The objective of the mid-term evaluation was to determine the feasibility of an econometric evaluation of the PIA, to identify and select from amongst all types of quasi-experimental methods the most relevant one to analyse the statistical quality of the chosen sample, to choose relevant indicators and potential control groups with a processing group to experiment quantitative and qualitative methods based on data reported from a sample of 50 completed projects, and validate the reliability of the tools developed.	The quantitative-qualitative ex-post evaluation method was expected to quantify the economic and environmental impacts of the aid on beneficiary companies, as an alternative to counterfactual statistical modelling.

Case study	Approach / type of use	Comment or quote
[IE] Better Energy Homes	The billing data analysis was compared with the simplified engineering calculations.	Differences in the results might be due to the effects of behavioural changes (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 estimates and the ex-post billing analysis, but not all.
[IT] White Certificates Scheme	Different methods are used to take into account the differences in the complexity of the projects. But there is no comparison between methods.	
[LT] Renovation programmes with EU funding	Comparison of scaled savings based on energy audits and building energy certification with metered savings.	The comparison emphasised the importance to monitor complementary parameters (temperature, moisture content and CO_2 concentration).
[NL] Multi-year agreements in the industry	Change in 2010 from a top-down approach (monitoring of energy efficiency indicators based on energy use per unit of production) to a bottom-up approach (monitoring of the actions implemented and energy savings per project).	The bottom-up approach is assumed to give a better view of the companies' efforts to save energy (actions implemented). Whereas the top-down approach was assumed to give a view of the "actual" energy efficiency trends (as based on data of actual energy use). Research about the comparison of the two methods showed that the choice of the monitoring method affects the key message to policy makers.
[NL] Subsidy scheme for housing corporations	The evaluation aimed at assessing "actual" energy savings to compare with the deemed savings. The evaluation used an approach similar to stock modelling (due to data issues). Further research could make an analysis based on metered energy consumption before and after renovation works.	The comparison of theoretical energy use (energy certificates) and actual energy use (energy bills) showed a prebound effect (lower actual gas consumption vs. theoretical consumption), this effect being larger when the energy performance of the dwelling is lower. The complementary research showed that using average values from the dwelling stock also introduced differences with the energy savings observed for the dwellings that were renovated.
[NL] Fiscal incentives for cars	Comparison of the standardised tests done in laboratories with field measurements (differences between standardised driving cycles and actual driving conditions and behaviours).	The ex-ante evaluation was based on paper data (values derived under specified test conditions) and the ex-post evaluation was based on actual emission measurements. This resulted in a difference of 10-20 %.
[UK] Supplier Obligations	Deemed savings are regularly revised based on further studies or analysis, especially based on the National Energy Efficiency Data-framework (NEED).	NEED has been used to understand the reduction in consumption for households installing specific energy efficiency actions using metered energy consumption data rather than modelled energy use.

Case study	Approach / type of use	Comment or quote
[UK] Warm Front	The particularities of the Warm Front scheme (with the objective of alleviating fuel poverty) led to the reporting of two different results: 1) high theoretical energy savings (when assuming normalised heating behaviours and considering the improvements of dwellings' energy efficiency); 2) no significant energy savings, when taking into account comfort and underperformance factors (i.e. estimates based on changes in energy consumption as observed on energy bills).	The case of Warm Front provides two different approaches used to question or refine the savings estimates based on simplified engineering calculations. For more details about the comparison and the analyses of changes in energy consumption as done in the research project about health impacts, see the EPATEE case study about the Warm Front scheme.
[US] Weatherization Assistance Program	The latest official evaluation of WAP included the comparison of several methods, particularly about the methods used to normalise energy consumption according to weather conditions. WAP was also the subject of other evaluation research by academics. "The national evaluation was attempting to develop an overall estimate of program savings by collecting and reporting information from all parts of the country, covering all types of fuels, and all types of buildings. The national evaluation showed that there was considerable heterogeneity in the savings estimates among regions, states, and agencies. (…) The national evaluation also showed that there was considerable heterogeneity by pre-weatherization usage, building type, main heating fuel, and source of funding. (…) The Fowlie study examined a quite different question. They focused their research on the service territory of one agency in Michigan. And, they asked the question … "What would happen if you attempted to deliver WAP program services to EVERY income-eligible household in that service territory." Their study was quite good at answering the question that they designed their study to answer. However, their study said next to nothing about the national WAP program."	The different academic studies showed results different from the official evaluation, which raised debates. The main conclusion from this experience is that different methods or evaluation approaches might bring different views on the programme and its results. One should therefore be aware of what each method is meant for, when considering their results. See more details in the "Focus on" section of the EPATEE case study on WAP. "There are different ways to conduct studies. It is important to look at these different ways, acknowledging that they might bring different insights about the program. Then when making comparisons, it is essential to be careful to compare things that are comparable."

6.4 Uncertainties and cautions when analysing data

Table 19. Sources of uncertainties identified in the EPATEE case studies.

Case study	Sources of uncertainties
[AT] Environmental Support scheme	Reported energy savings are based on information provided by the applicants and are ex-ante.

Case study	Sources of uncertainties
[AT] City Energy Efficiency Programmes of Vienna	The top-down analysis shows the overall achievement of the target but does not show the contribution of the SEP or any other causality. However framework conditions (climate, population growth, financial crisis of 2008/2009) are taken into account in the analysis. To reduce this uncertainty a bottom-up calculation of implemented energy efficiency measures was done. The uncertainties of this bottom-up methodology (mostly deemed savings and engineering estimates) are connected with the underlying ex-ante approach and the need to make certain assumptions on consumer behaviour (incl. room temperature), climate conditions and the technically optimal installation of technologies. In addition not all actions implemented in the framework of the SEP can be evaluated bottom-up.
[BE] Primes Energie	differences between reference values and actual characteristics of participants' dwellings (for the baseline situation); differences between the assumptions on heating behaviour and the actual behaviour of the participants (for example due to rebound effect); differences between the assumed (deemed) and actual performance of the actions installed; errors in the data in the files submitted to get grants (when outliers are identified, then they are not taken into account in the calculations).
[CR] Energy renovation of public sector buildings	differences between ex-ante estimates (deemed savings) and actual (metered) energy savings: currently the consumption and action databases are not connected, but the objective is to enable this connection in the future, to compare ex-ante estimates with actual energy consumption monitored ex-post, which would drastically improve calculation's accuracy; use of values specific to the projects implemented is recommended, but use of default values is also allowed; no correction factor for rebound effect.
[CR] Individual heat metering in multifamily buildings	Differences between ex-ante calculation (deemed estimates) and actual energy savings (based on billing analysis). Use of default values. There is no systematic ex-post monitoring. Ex-post results are available from the sample investigated in the ex-post evaluations and needed to be extrapolated to the whole buildings where heat cost allocators were installed. This extrapolation includes statistical uncertainties. The different evaluation studies done showed different results.
[DK] EEO scheme	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 (standard values). Use of various methods to consolidate the conclusions about net impacts and additionality. In the 2015 evaluation, a question in the survey of participants was focused on the energy savings: only 4% of 321 surveyed participants said they observed lower energy savings than estimated (but 55% said it was too early to know or they don't know).
[FI] Energy Efficiency Agreement for Industries	The accuracy of the savings calculation for reported individual actions corresponds to the accuracy that may be achieved in normal field work and the calculations are typically carried out by an external consultant on behalf of the participating party (e.g. energy auditor). Some of the initial data are based on design/manufacturer data of technical systems or estimates, since measurements are often not possible or at least too costly. Savings calculations of individual actions dependent on outside temperatures are made using normalized energy consumption data. Measures have been taken to reduce uncertainty by providing guidelines, training and support: A guideline has been issued for agreement participants and their service providers for estimating the energy savings impact of reported energy efficiency actions. At the beginning of annual reporting, briefing sessions have been organised for the participants and their service providers. Since 2014 this has been implemented mainly as webinars. Companies are supported via a designated email service.

Case study	Sources of uncertainties
[FI] Voluntary energy audits for municipalities	The accuracy of the savings calculation for reported individual actions corresponds to the accuracy that may be achieved in normal field work and the calculations are typically carried out by an external consultant on behalf of the participating party (e.g. energy auditor). Some of the initial data are design data of technical systems or estimates, since measurements are often not possible or at least too costly. Savings calculations of individual actions dependent on outside temperatures are made using normalised energy consumption data.
[FR] Voluntary agreement for freight companies	The main uncertainties about estimation of CO_2e emissions reduction come from the internal data collection, measurement & reporting procedures of transport operators (especially for operators that assessed their emissions from default values). Participating companies are encouraged to have their data independently verified but this is not mandatory (excepted for those applying for the scheme label). The data collected for the ex-post evaluation from an online survey may include additional uncertainties/errors/bias due to questions formulation, level of knowledge of the surveyed person and their availability when the survey was performed. Some uncertainties are due to extrapolation of some data (see section 6.2).
[FR] "Future Investments" programme	Uncertainties in project costs estimation may be due to early estimation (often many years before the projects start) and due to changes in the project (withdrawal of a partner, technical barriers causing technological changes….). Uncertainties may also be due to the very innovative level of some technologies whose development costs are very difficult to assess, and not in line with market price. This led to uncertainties in the development of the scenarios that are used to set the refundable conditions in case of commercial success. The competitive context is also difficult to assess for innovative technologies (difficult to define the pre-existing or future market for technologies not yet available on the market). These multiple changes may also lead to errors in the monitoring database which is not systematically updated. Uncertainties have also been observed when assessing the environmental effects of the technologies compared to a reference situation (in case of projects for environment protection).
[DE] Energy Efficiency Networks Initiative	The savings estimated based on the targets defined by each network are likely underestimating the actual savings because it is safer for network participants to set the ex-ante target lower than possible, to not miss it. In a second step of evaluation, after predefined network operation time, companies are evaluated for network achievement. Actual savings are then compared with network savings targets. Preliminary results of a first sample group point towards an average target achievement of 110% of the predefined network target. The survey method bears uncertainties because no physical on-site measurements are performed. Hence, submitted savings cannot be easily verified. To limit this uncertainty, a verification is performed for randomly selected 10 percent of the participants. Of these companies, documentation like audit reports or project documentation is requested. However, the verification occurs document based. No on-site measurements are performed.
[DE] Energy Efficiency Fund	Differences in calculation methodologies used by energy auditors (for complex actions). Uncertainties of values from product data sheets (for simple projects, no ex-post verification). Targets defined in emissions reductions. Calculation based on assumed constant emissions factors (that can change over time in practice, for example due to changes in the national electricity mix). Effects (e.g. free-rider effect) calculated based on surveys (risk of bias in the answers).
[IE] Better Energy Homes	Differences between ex-ante estimates (deemed savings) and actual energy savings (as evaluated with the billing analysis using a difference-in-differences method). The confidence interval for the ex-post evaluation (± 603 kWh/year) represents the uncertainty range that can be assessed thanks to the statistical approach. However it does not capture all the sources of uncertainty (for ex., bias in sampling and matching).
[IT] White Certificates Scheme	Errors in the calculations and reporting of the energy savings, which are monitored by GSE by documentation controls and on-site inspections. Uncertainties related to the use of engineering calculations or deemed savings (until 2017) and in a new standardised method introduced by Decree 11 January 2017, which is based on measurements on samples of interventions. Uncertainties related to monitoring plans related to energy processes in which a large number of

Case study	Sources of uncertainties
	variables is involved.
[LT] Renovation programmes with EU funding	Energy savings are calculated based on the data from the energy certification of the building before and after renovation works and energy audit performed before renovation. The main uncertainties come from the fact that no measured data on energy consumption is used for energy audits for at least 3 years both, before and after renovation, to make data reliable. Besides, measurements of microclimate indicators (temperature, moisture content and CO_2 concentration) were not defined before and after renovation for defining living standards inside renovated buildings. Energy savings assessment in a sample of about 50 renovated buildings based on measured heating data showed actual energy savings of about 30% vs. reported energy savings usually around 50%.
[NL] Multi-year agreements in the industry	A major source of uncertainty about the energy savings reported relates to the fact that the savings are self-assessed/monitored by the companies themselves. Although monitoring reports submitted by the companies are checked for completeness and correctness by external consultants, the check is more of a probability check.
[NL] Subsidy scheme for housing corporations	The energy savings were calculated with estimated gas savings per label step according to data on real energy use for a large sample of houses (other than the dwellings renovated within the scheme). The average savings per label step have been calculated as a difference between the energy use of two distinct groups of dwellings having distinct energy labels (for ex. E and C). This difference in energy use may reflect the differences in energy efficiency of the building envelope and heating systems, but also the differences in the energy behaviours of the occupants. However the large size of the sample makes possible to have average values per label that are less sensitive to the variations in energy behaviours or to the specificities of each building. The complementary research done in 2016 provided data not only about the subsidized energy saving actions that reduce the gas use like insulation or condensing boilers, but also actions such as connection of houses to a district heating network or installation of solar PV. The estimated gas savings per label step don't take into account these actions.
[NL] Fiscal incentives for cars	"The effectiveness of the measure in reality (emission of new cars in real life conditions) is lower than the effectiveness of the measure based on paper values (emission of new cars under standard test conditions)." The uncertainty related to the evaluation of the impact of the tax measure on the sales of cars (based on stock modelling) are mainly due to the uncertainty in the data from CBS (the Dutch national statistics bureau), that are key data for modelling the changes in the stock of cars. The uncertainty in the data for energy savings is in the estimate of total sold number of cars with the purchase tax advantage, and the total amount of km's driven by these new cars, the total amount of avoided CO_2 emission of these cars and the related total avoided energy use (expressed as primary energy avoided).
[Nordic Countries] Nordsyn	Authors of the Effect study highlighted that the approach to evaluate the impacts ("lost" energy savings) of non-compliant products only capture the "visible" part of market surveillance impacts. An ideal approach would be to compare two regions/countries, one with and one without market surveillance. However, all EU Member States are required to implement market surveillance. When considering the market surveillance practices, the Effect study had to assume that the same product is not tested multiple times by different countries, that the markets are similar in all Nordic countries, and that all countries would act on all test results (i.e. remove "instantly" the products detected as non-compliant in any of the Nordic countries from their national market). A major source of uncertainties is due to the extrapolation from results for Denmark to the other Nordic countries, proportionally to the GDP of each country. This means an implicit assumption that the sales of products would be proportional to the GDP, with similar types of products in all Nordic countries. Due to choices made to define the baseline and the different sampling methods, several assumptions were needed. The evaluators chose to make conservative assumptions, which leads to underestimate the "lost" energy savings. This choice was made because it was not possible to assess the uncertainties, therefore using conservative assumptions was a way to provide results as reliable as possible. Also, they mention that more precise data on the lifespan, cost of market surveillance, sales, and electricity prices could improve the accuracy of the calculation. Monitoring only the non-compliance to standards shows only a limited part of the lost energy

Case study	Sources of uncertainties
	savings potential, since there are more ways a product can use more energy that is effective; e.g. the way the product is used, or lack in technical documentation.
[UK] Supplier Obligations	errors in the calculations and reporting of the energy savings (addressed by random checks); uncertainties related to the use of deemed savings (e.g., differences between estimated and observed energy consumption) and to the in-use factors (defined to correct part of these differences, particularly for performance gaps and rebound effects). An evaluation by the National Audit Office (2016) refers to analysis by the Department of Energy & Climate Change (now Department for Business, Energy & Industrial Strategy) carried out in 2015, which evaluated ECO, CERT and CESP on the same basis. This analysis found that CO_2 savings attributable to CERT and CESP were 50% lower than it originally calculated, largely a result of using different (and more realistic) estimates for the energy savings from specific actions. However, similar analysis is not available for the years prior to 2008. Actual savings for the years 1994-2008 may therefore be lower than reported savings.
[UK] Warm Front	Uncertainties related to the simplified engineering calculations: overestimation of the heating consumption before intervention (for ex., due to restriction behaviours), also called “prebound effect; underestimation of the heating consumption after intervention (for ex., due to comfort taking/rebound effect, defaults in the installation of the actions, etc.). About the study done within the research on health impacts, the main sources of uncertainty were related to data limitations (consumption mostly metered on short periods of 2 to 4 weeks, indoor temperature not monitored in all rooms).
[US] Auctions for capacity markets	errors in the calculations and reporting of the energy savings from deemed savings (addressed by random checks); uncertainties related to the use of deemed savings in standard projects (e.g., differences between estimated and observed energy consumption); uncertainties around the baseline used in customised projects with tailored M&V plan.
[US] Weatherization Assistance Program	Care was taken to limit uncertainties due to sampling, as well as selection biases when calculating the total annual energy savings and the average annual per-household savings. Regarding the latter, the ORNL (Oak Ridge National Laboratory) aggregate model was applied as it identifies baseload consumption allowing to estimate uncertainties in parameters and calculation of values with a statistical basis. Sensitivity Analysis was foreseen to be undertaken for the PY 2010 retrospective evaluation, to observe how yearly estimates of energy and costs savings may alter due to changes in key driving factors, such as changing demographics in the houses, loss of housing stock, volatile fuel prices, technology evolution, and weather conditions. Sensitivity analysis was also foreseen for determining the impact of key assumptions used in the calculation of Savings-to-Investments Ratio (SIR). This approach would allow the assessment of the influence on SIR of uncertainty in key assumptions (e.g. real discount rate, action lifetime, monetary value of non-energy impacts), from probability distributions. However, this analysis could finally not be completed.

7 | Evaluation practices: methods and feedback about evaluating net impacts

The main conclusion from this part is:

MESSAGE 09: Analysing causality is a challenge, but essential to assess efficiency of policies.

The table below gathers the feedback collected about evaluating net impacts or assessing additionality. This shows diversity in the way to define or handle the concept of net or additional impacts, as well as in the methods used. Difficulties were also often reported, sometimes impeding the use of the method initially planned. For more details about the methods to evaluate net impacts (and practical examples), see the EPATEE topical case study on evaluating net impacts.

Table 20. Methods and feedbacks about the evaluation of net impacts.

Case study	How causality was defined/evaluated	Comment / quotes
[AT] Environmental Support scheme	No ex-post causality assessment. Savings are calculated vs. "actual before", but performance of projects needs to be additional (additionality criteria). The verification includes a focus on the additionality of the projects (additionality criteria: performance > regulations, and payback time > 3 years).	Emphasis put on the additionality of the performance of the projects and investments (monitoring of marginal costs), not on the additionality of savings. Performance-based incentive: the higher the environmental/energy performance, the higher the subsidy. The evaluation does not differentiate between gross and net savings. This would require additional evaluation efforts. In practice, it is very difficult to assess free-riders or additionality. Surveys might include bias, for example as participants tend not to tell that they would have done the action anyway. In general there is a lack of data to provide evidence on this issue.
[AT] City Energy Efficiency Programmes of Vienna	Additionality of savings is achieved by using the official Austrian ESD/EED methods that have been publicly available since 2008 and that are updated regularly. Ex-ante approach about additionality (no ex-post evaluation of the causality).	This means in particular that the savings of the programme are additional compared to minimum energy performance requirements enforced by EU and Austrian regulations, and that double counting with Federal measures is avoided.
[BE] Primes Energie	Savings are calculated vs. "stock average", but performance of projects needs to be additional (additionality criteria). No ex-post causality assessment, but qualitative email survey about the free-rider effect of another measure in the same portfolio (EcoPack scheme).	Eligibility criteria on actions ensure that they go beyond minimum legal requirements (performance additionality). "Additionality is also a very challenging issue. Our approach is to attribute the energy savings to the measure for which a direct link can be established between the measure and the acting decision. We do acknowledge that the acting decision results from a combination of factors. But this approach makes possible to avoid risks of double counting. However then, things get more complicated if the evaluation has to take into account the additionality rules set by the EED (Annex V). For example, when a boiler is replaced: would it have been replaced anyway (in the absence of the grant)? In this case, it may be relevant to use as baseline the EcoDesign requirements for boilers. But if the boiler would not have been replaced, or replaced a few years later, then it would be more relevant to define the baseline situation as the efficiency of the boiler that was replaced. In practice, this is too difficult (and costly)

Case study	How causality was defined/evaluated	Comment / quotes
		to assess and to implement in the evaluation. More generally, it is very difficult to distinguish the effects of a measure from all the other factors that affect acting decisions. Particularly for measures that are already implemented for a long time. This is the case for the Primes Energie scheme that started in 2004. How to know what would have happened now, if the Primes Energie scheme had never existed? Meanwhile, the experience acquired with monitoring and evaluating the scheme over many years makes that we have some qualitative understanding about how it may affect acting decisions. One lesson learnt is for example that the incentive should be attractive, not only financially attractive but also in terms of simplicity of use. (…) another lesson learnt is that trying to limit free-rider effects may lead to unexpected negative effects (for ex., if the decrease in the applications is stronger for households who would be the most in need)."
[DK] EEO scheme	Survey of participants done in each ex-post evaluation (but raising debates). In 2012, changes in energy consumption of a control and participants group (about 160 households each) were compared over 24 months, showing that the net effect for the participants group would be about 56% of the energy savings achieved (caution: small samples, not meant to be representative). In 2015, econometric analyses (panel data regression and co-integration) were used to assess net effects at sector level. However results were conclusive only for industry (74% of additional actions). The data available in other sectors were not disaggregated enough to make possible to distinguish the effects of the scheme from effects of other factors (e.g., energy prices).	"Frequently, evaluations aim to determine the net impact of a policy. Finding a relevant control group to use quasi-experimental approaches (statistical comparisons of a participants and a control group) is often not possible and can be expensive. So, in many cases, the only option left is to use surveys. Therefore, evaluators at times have to rely on surveys where the respondents are self-reporting. Typically, the evaluation takes place some time after the activity and the end-users may not remember exactly what took place, why they chose to participate or not, what the cost was, and more importantly would they have acted differently if they had not participated. And finding the relevant person to talk to can be difficult, especially when it comes to non-residential consumers." "Another lesson learnt is that additionality is a difficult topic for most to grasp. It is not necessarily cost-effective to strive for 100% additionality – on the contrary. And this fact is difficult to communicate." See more details in the example included in the EPATEE topical case study about net energy savings.
[FR] Voluntary agreement for freight companies	The results were assessed in terms of gross impacts (before/after comparison). The survey of participants was used to analyse their motivations to join the scheme and implement actions. But this was not used to assess net impacts.	Based on the survey, it was estimated that 239 out of 290 companies would have implemented the actions without the scheme. However the reliability of the answers from the respondents is unknown (particularly about social desirability bias).
[FR] "Future Investments" programme	Ex-ante: Net impacts evaluated in terms of additional environmental or energy performance compared to a counterfactual that should represent the level that would be achieved without the aid. Ex-post: two different methods tested: 1) Difference-in-Differences (DiD) method, and 2) survey of participants.	About ex-ante evaluation: the definition of this baseline is made according to European Commission's Guidelines on State aid for environmental protection and energy 2014-2020. About ex-post evaluation: the samples available for the DiD method did not make possible to obtain statistically significant result (in terms of causality assessment).

Case study	How causality was defined/evaluated	Comment / quotes
[DE] Energy Efficiency Networks Initiative	No causality assessment. Additionality taken into account in the baseline.	Additionality in the sense of the EED is addressed using the baseline of minimum standards to avoid double counting with existing minimum requirement policies.
[DE] Energy Efficiency Fund	Free-rider effects determined based on ex-post surveys. Moreover, the whole Energy Efficiency Fund evaluation considers interaction effects between the different measures.	There might be risk of biased answers in the surveys to assess free-rider effects, and also issue of sample size. "In the exhaust heat evaluation, the survey sample size was too small to be used for a reliable calculation of free-rider effects."
[IE] Better Energy Homes	Default value (18%) for free-rider effects (based on the evaluation of the Energy Efficiency Commitment in UK) were taken into account in the cost-benefit analysis. The billing analysis used a difference-in-differences method (comparing changes in participants and control group).	Free-rider effects not taken into account in the energy savings reported for the regular monitoring.
[IT] White Certificates Scheme	Causality is handled beforehand by setting additionality requirements, both for projects to be eligible and for defining the baseline used to calculate additional energy savings.	The additionality requirements have been strengthened over time. Moreover since 2013, projects received a national public aid (e.g. tax credit) are no longer eligible for white certificates (to avoid double counting between policies).
[NL] Multi-year agreements in the industry	Energy savings in the annual reports are gross energy savings (no causality assessment).	"Whilst the NL Agency has put in place a monitoring system for the LTA scheme, a major weakness is the fact that there are no benchmarks or counterfactual differences, this makes the ex-post evaluation difficult." "the 2008-2012 ex-post evaluation carried out by external parties, under the leadership of Ecorys, was challenged by lack of approach on how to take into account endogenous technology development in the assessment methodology as well as difficulties in assessing where in the chain the efficiency is realised, i.e. disentangling and assigning the energy efficiency to the different parts of the chain. For example, some energy savings are the result of regular replacement of equipment, since newer equipment is almost always more efficient than the old equipment. This is not a policy effect but an "autonomous savings" effect."
[NL] Subsidy scheme for housing corporations	Net impacts could not be evaluated.	The evaluator mentioned that if before the introduction of the subsidy scheme, the policymakers would have made an inventory about the energy actions that housing corporations already do (business-as-usual), then a baseline could have been set. That did however not happen. So in the evaluation only gross results could be evaluated, no information is available about the results in case of absence of the subsidy scheme.
[NL] Fiscal incentives for cars	Net energy savings assessed compared to the average trend observed in EU countries.	This approach can be considered a type of control group, where the benchmarking with EU countries would represent the control group.
[Nordic Countries] Nordsyn	The evaluation provides an assessment of non-compliance rates per type of appliance. It was not possible to compare with non-compliance rates in other countries.	As the evaluation deals with a regulation, free-rider effect does not apply to this case. The net impact could be evaluated as the additional impact from extra efforts of market surveillance compared to the market surveillance in other countries.

Case study	How causality was defined/evaluated	Comment / quotes
[UK] Supplier Obligations	Estimates used to revise the deemed savings are based on observed savings, so they are savings after comfort taking (direct rebound effect). To do this, intervention and comparator groups are created – with the intervention group containing properties which have received the energy efficiency action being considered (and no other action), and the comparator group containing similar properties that have not had a recorded energy efficiency action installed at any point. However this approach does not inform about causality or free-rider effects.	"Counterfactuals are always difficult to define as there are no areas in the UK that have not been treated under the Supplier Obligation."
[UK] Warm Front	The evaluations did not include an explicit causality assessment, assuming implicitly that all the actions getting a Warm Front grant would not have been installed in the absence of the scheme. The causality issue in the case of the Warm Front scheme was not to be analysed in terms of "usual" additionality or free-rider effects (participants who would have done the action anyway) but in terms of targeting, as the priority objective of the scheme was to alleviate fuel poverty: did the scheme provide grants to households that were in situation or at risk of fuel poverty?	The assumption of absence of free-rider effect seems to be confirmed for most of the participants surveyed in the health impact evaluation and in the process evaluation. But a few surveyed participants felt they would have been able to afford the actions without help from Warm Front or other sources. Similarly, the results from the survey supports the assumption that the spill-over effects at the level of the participants would be limited. At the opposite, the high level of satisfaction of the participants may have led to significant spill-over between households (word-of-mouth), with several surveyed participants reporting that they recommended the scheme and installers. About the targeting of the scheme, several adaptations of the scheme were made over time, particularly taking into account the recommendations from the reviews done by the National Audit Office.
[US] Auctions for capacity markets	Causality assessment is not required for this scheme. Focus is on the reliability of the resources (M&V plans). Additionality is handled with the rules for defining baselines.	The baseline is defined depending on when the equipment is replaced: For equipment replaced before end of its operational lifetime, the baseline is the average load of electrical equipment being replaced ("before actual"); For equipment replaced after end of its operational lifetime, or new installation, the baseline is the efficiency standard or industry standard practice ("minimum performance standards").
[US] Weatherization Assistance Program	The evaluation studies assumed there was no free-rider effect since most of homes in the WAP are low-income with documented problems in paying everyday bills. The use of a comparison group showed a small reduction in weather normalized usage during the analysis period. That indicates that there might have been "some" actions taken by the comparison group, but to a limited extent. That also shows the importance of using a comparison group in the evaluation.	"About attribution methodology, the idea behind the attribution analysis was to try to assess the contribution of each funding source to the outcomes of the delivery of WAP program services. However, the data collected from the individual agencies did not furnish enough detail on how the funding from each source was used, thereby limiting our ability to attribute savings to specific funding sources."

8 | Beyond the evaluation of energy savings

The priority of the EPATEE project is to look at the evaluation of energy savings from energy efficiency policies and programmes. Therefore the analyses in the case studies were focused primarily on reviewing the data on energy savings and the methods used to evaluate them.

Preliminary interviews and online surveys done by the project to assess stakeholders' needs and priorities highlighted their high interest in the evaluation of other indicators and impacts than energy savings. Therefore the case studies were also used to review what types of other indicators or impacts were evaluated together with the energy savings. However, the EPATEE case studies did not intend to analyse the methods used to assess these other indicators or impacts. Moreover, the review done was based on the evaluation reports including data about energy savings and on interviews with stakeholders involved in these evaluations. Other evaluation studies might have been done that were not identified by the EPATEE case studies.

Section 8.1 presents the type of indicators used to assess cost-effectiveness and cost-efficiency (either at the project or policy level). Section 8.2 presents an overview of the information found about the types of impacts other than energy savings. Section 8.3 presents what other aspects (process evaluation, customer journey, participants' satisfaction, market transformation, etc.) have been evaluated, together or in parallel of the evaluation of energy savings.

8.1 Cost-effectiveness and cost-efficiency

Table 21. Indicators used to assess cost-effectiveness and cost-efficiency (either at project or policy level).

Case study	Indicator about cost-effectiveness	Indicator about cost-efficiency
[AT] Environmental Support scheme	Total environmentally relevant excess investments triggered (marginal costs to achieve energy or environmental performance beyond regulations)	Public efficiency indicator (amounts of subsidies in euros per tCO_2 avoided over the lifetime of the projects funded)
[BE] Primes Energie	Public cost per energy saved (c€/kWh saved)	
[CR] Energy renovation of public sector buildings	At the action level: cost effectiveness of implemented actions are compared with cost optimal actions that are obtained by calculating costs and benefits for several predefined scenarios for each building	
[CR] Individual heat metering in multifamily buildings	At project level: net present value of investment and energy savings	
[DK] EEO scheme	Cost-effectiveness of the obligated parties' programmes: costs incurred by the obligated parties per reported kWh saved.	Socio-economic cost-effectiveness: socio-economic net value of additional energy savings projects over the lifetime of the actions.
[FI] Energy Efficiency Agreement for Industries	Direct payback time, cost saving (on energy bills) and investment cost are monitored (based on data reported by the participants) A very important indicator for the participants themselves is the cost savings achieved through participation.	
[FI] Voluntary energy audits for municipalities	Same as for voluntary agreements	

Case study	Indicator about cost-effectiveness	Indicator about cost-efficiency
[FR] Voluntary agreement for freight companies	Total efficiency: total expenses (in euros) divided by the CO_2 emissions avoided (in tCO_2)	Public efficiency: total public expenses (in euros) divided by the CO_2 emissions avoided (in tCO_2)
[FR] "Future Investments" programme	At project level: Net Present Value.	Leverage effect, fiscal multiplier of public expenditure (BFTB: "Bang for the Buck")
[DE] Energy Efficiency Fund	Leverage effect in terms of total investment triggered by funding (Euros of total investment per Euro spent for funding)	Energy savings or emissions reductions per 1000 Euros spent (funding amount + administrative cost)
[IE] Better Energy Homes	Net Present Value (NPV) of the scheme (taking into account 3 different views: government, participants and the whole society); Cost-effectiveness indicators (in euros/kWh or tCO_2 saved) based on a Cost-Benefit Analysis	
[IT] White Certificates Scheme		Public expense for an additional electricity kWh saved (monitored until 2011)
[LT] Renovation programmes with EU funding	Monitoring of the average renovation cost per m^2	
[NL] Multi-year agreements in the industry	Implementation costs for government and industry in relation to the benefits	
[NL] Subsidy scheme for housing corporations	Public budget (amount of subsidies) divided by the reduction in CO_2 emissions (euros/tCO_2)	
[Nordic Countries] Nordsyn	Cost-benefit ratio of market surveillance calculated as return on investments (costs of market surveillance activities vs. savings on electricity bills by removing non-compliant products from the market)	
[UK] Supplier Obligations	Net Present Value (ex-ante assessment); Value to society of lower energy bills in low income households (ex-ante assessment)	
[UK] Warm Front	Value for money in comparison with the general market	
[US] Weatherization Assistance Program	Savings to Investment Ratio (SIR): lifetime energy bill savings divided by the costs, both for actions done in a given year. The programme is deemed cost-effective if SIR > 1.	

Complementary qualitative feedback from some of the cases is also worth to consider, and shows the importance to look at impacts other than energy savings, as well as at qualitative aspects:

- From the case on Primes Energy (Belgium – Wallonia):

> "This experience showed that it is very important to know the effectiveness of the measures. It is indeed essential for the discussions about budget. Effectiveness is first analysed in terms of public cost per energy saved (c€/kWh saved). But other criteria are also taken into account for decision making, like the policy priorities set by the government."

"When energy efficiency is not the priority objective of the measure, then it may also be difficult to assess its effectiveness in terms of public cost per energy saved. Sometimes only some part of the public cost should be attributed to the energy efficiency objective. Such assessment requires a good understanding of the measure and its policy theory to use the right assumptions."

- From the case on Better Energy Homes (Ireland):

  About the CBA done in 2010, the evaluators pointed that, despite all the efforts done to use the best data and methods available, a number of factors could affect the assumptions made. Sensitivity analyses were therefore conducted to deal with this uncertainty and to investigate the impact of varying key assumptions such as energy prices over time and taking into account (or not) externalities (reduction in CO_2 emissions and in other air pollutants), thus producing a range of possible outcomes. The results of the different scenarios tested showed the major importance of the assumptions done on the trends in energy prices over the action lifetime.

  The evaluators of the CBA also highlighted that all externalities cannot be monetised and/or included in a CBA. They warned that the results of a CBA do not provide the only basis for policy assessment and selection. These analyses need to be considered together with qualitative analysis of other costs and benefits not monetised in these calculations. About the BEH scheme, these might include for example potential job creation, increased comfort and property value, and potential health improvements.

- From the case on the Warm Front scheme (England):

  "This research [early research project about health impacts] showed the usefulness to make evaluation with a broad scope. Evaluations are often focused on a limited set of indicators that might not be the most relevant. This is usually because funding for evaluation is often limited.

  Therefore, evaluation is ordered with a restricted approach and the main objective to assess whether the policy worked or not. However, this is only one part of what evaluation should be. Most of the time, it would be more important to understand WHY the policy worked or did not work."

- From the case on the Weatherization Assistance Program (US):

  "It should be noted that the cost-effectiveness can be assessed according to two scopes, one taking into account the benefits in terms of energy savings only, and the other taking into account also the health benefits. This broader scope gets a higher attention in recent years, as health benefits prove to be important for the program."

8.2 Impacts other than energy savings

Table 22. Overview of the types of other impacts for which information was found.

Country	Name of the measure	CO2 emissions	job impacts	affordability of energy / impact on energy prices	distributive / sectoral / geographical effects	leverage effect (on direct investments)	Emissions of other pollutants	impacts on comfort (buildings)	GDP and other market impacts	impacts on public budget	Health impacts	Water savings / impacts
Austria	Environmental Support' (Umweltförderung im Inland)	X	X		X	X			X			X
Belgium (Wa	Primes Energie (grants for energy renovation)	X	X		X				X			
Croatia	Energy renovation of public sector buildings programme	X	X		X	X			X	X		
Croatia	Individual heat metering in multifamily buildings	X		X								
Denmark	EEO scheme	X		X	X		X		X			
Finland	Energy Efficiency Agreement for Industries	X				X						
Finland	Voluntary audits for municipalities	X										X
France	"Future Investments" programme	X	X			X						
France	CO2 target programme	X					X					
Germany	Energy Efficiency Fund	X				X						
Germany	Energy Efficiency Networks Initiative	X										
Ireland	Better Energy Homes	X	X				X	X		X		
Italy	White Certificates Scheme		X	X	X							
Lithuania	Renovation programmes with EU funding			X				X				
Netherlands	Subsidy scheme for housing corportations in Amsterdam	X		X				X				
Netherlands	Purchase tax reduction for efficient passenger cars	X								X		
UK	Supplier Obligations	X		X			X	X			X	
UK	Warm Front	X	X					X			X	
US	Weatherization Assistance Project	X	X				X				X	
	N = 20	17	8	6	5	5	5	5	4	3	3	2

Quotes from some of the interviews done for the EPATEE case studies also show the importance of impacts other than energy savings in the way policy measures can be appreciated by stakeholders:

"The main motivator for introducing heat cost allocators was the improvement of thermal comfort in our dwellings, while the reduction of costs was only a secondary motive."

Quote from the case on **individual heat metering in Croatia**.

"A fundamental reason is that saving energy is not a major driver in decision making on investments which are principally driven by reasons related to production, safety, health or environment."

Experience feedback from the **Finnish Energy Efficiency Agreements in Industry**.

"According to the survey of participants, the comfort improvement was the first tangible impact felt by households following an upgrade."

Experience feedback from **Better Energy Homes (Ireland)**.

"The quantification of energy savings makes sense, but should not be the only focus. Other appropriate indicators showing the success to specific measures have to be found and used."

Quote from the case on the **municipal energy efficiency programme of the City of Vienna**

"More generally, the evaluation confirmed that health impacts were larger than impacts on fuel consumption. The main improvements for the participants could indeed be found in terms of higher indoor temperature and better comfort. (…) This research showed the usefulness to make evaluation with a broad scope. Evaluations are often focused on a limited set of indicators that might not be the most relevant. This is usually because funding for evaluation is often limited."

Quote from the case on **Warm Front (England)**.

"Non-Energy Benefits (NEBs) have always been an important topic for WAP. The methodology tested during the last national evaluations is now used by the states for their own evaluation, which should bring more data and evidence about NEBs. (…) NEBs are critical outcomes of WAP. We need to continue the development and improvement of approaches to quantify NEBs. There is still room for improvements in this field."

Quote from the case on the **Weatherization Assistance Program (US)**.

8.3 Other aspects evaluated

Table 23. Overview of other aspects evaluated, as found in the EPATEE case studies.

Case study	Process evaluation	Customer journey / participants satisfaction	Market transformation	Others / comments
[AT] Environmental Support scheme	X	X		investigation of opportunities to improve the administration of the scheme, for example the average processing time (time between application and decision on eligibility)
[AT] City Energy Efficiency Programmes of Vienna	X			The structure of the programme was also looked at both with a view on overlaps of instruments as well as on data collection and other administrative issues.
[BE] Primes Energie	X			review of the scheme management (by the Court of Auditors); chronological analysis of the participation (to analyse how households react to changes in the scheme); needs in training for EE and building professionals (survey of companies)
[CR] Energy renovation of public sector buildings	X	X		Analysis of the financial and organizational issues is conducted through document analysis, in-field experience of the evaluators and through interviews with identified key stakeholders in the process.
[CR] Individual heat metering in multifamily buildings		X		Analysis of complaints from end-users, and reasons in case of higher heating bills (after the installation of heat allocators)
[DK] EEO scheme	X	X	X	Surveys of obligated parties, installers and end-users. Costs per type of actions (monitoring trends in obligated parties' costs); Qualitative assessment of the impact of the scheme on the energy efficiency markets (survey of contractors); Attempt to assess spill-over effects with an econometric analysis.
[FI] Energy Efficiency Agreement for Industries	X	X		Monitoring of the number of actions done using the ESCO model; Various indicators about the approach used by the participants (energy management, etc.); Satisfaction and feedback from the participants.
[FI] Voluntary energy audits for municipalities				Audited building volumes per building types; Number of actions recommended and recommendations implemented (per action type).
[FR] Voluntary agreement for freight companies	X	X		Online survey of participants, interviews with various stakeholders and case studies in a sample of regions. This enabled to assess the coherence and relevance of the scheme, and to make suggestions to improve it.

Case study	Process evaluation	Customer journey / participants satisfaction	Market transformation	Others / comments
[FR] "Future Investments" programme	X	X	X	The DiD method looked at direct impacts on participants' achievements of environmental objectives, investments in R&D and in environmental fields + indirect impacts on participants' business development (production, employment, etc.) and leverage effect. A survey was done to investigate impacts on the turnover of beneficiaries, jobs created or maintained in the investment and operational phases, CO_2 emission and other environmental impacts when relevant. In addition, the survey provided qualitative information on project management and the effects of aid on innovation, partnerships, collective learning, and commercial and technological opportunities.
[DE] Energy Efficiency Networks Initiative		X		"Each year, a survey is conducted among the network administrators and the participants" Monitoring of statistics about the networks (type, location, runtime, etc.)
[DE] Energy Efficiency Fund		X		Qualitative assessment from surveys with beneficiaries and administrators, with a particular focus on satisfaction with the provided information, processing times and support offered by the scheme.
[IE] Better Energy Homes		X		Survey of participants to better understand the reasons to take part in the scheme, and to assess their satisfaction with the scheme and actions done
[IT] White Certificates Scheme				Analysis of where the certificates are obtained (per sector and action type), and of market behaviours (e.g., capability of prioritising the most cost effective solutions)
[LT] Renovation programmes with EU funding	X	X		National Audit Report made in 2010 equivalent to a process evaluation (with recommendations to improve the programme) Survey to investigate satisfaction of flat owners and residents after renovation works. Barrier analysis and benchmarking of experiences in other countries.
[NL] Multi-year agreements in the industry	X			
[NL] Subsidy scheme for housing corporations		X		Comfort improvements and affordability of energy (ratio of energy expenses on total income) (assessed through survey of occupants) Housing corporations also mentioned that replacement of older gas installations improves safety and health.
[Nordic Countries] Nordsyn				The evaluation made suggestions about the strategy for appliance tests.
[UK] Supplier Obligations	X	X		Value of comfort taking (ex-ante); Process evaluation (suppliers' strategies, targeting of vulnerable customers, impact on the EE industry, role of public authorities).
[UK] Warm Front	X	X	X	Targeting and performance of the scheme monitored and assessed through various indicators and approaches; Satisfaction of the participants and comfort taking; Benefits or negative impacts on the supply chain (e.g., on turnover, jobs).

Case study	Process evaluation	Customer journey / participants satisfaction	Market transformation	Others / comments
[US] Auctions for capacity markets				For utility EE portfolios that bid into the ISO-NE forward capacity market, they are typically also subject to other types of evaluation than the peak savings verification. These evaluations are valuable for understanding how the design and implementation of EE schemes can be improved and forecasting market trends and system operational needs.
[US] Weatherization Assistance Program	X	X		Participant home (i.e., occupant, recipient) survey about budget issues, energy conservation behaviours, use of programmable thermostats and health issues. Case studies at state's level to analyse of how WAP services are delivered, and how weatherization staff and clients perceive service delivery. Special Technical Studies to evaluate performance of the WAP with respect to technical issues such as air sealing, duct sealing, etc. Weatherization Deferral Study to explore patterns across estimated deferral incidence rates and success with weatherization post-deferral and to highlight strengths and weaknesses within the deferral process Social Network Study to explore the potential for WAP recipients and staff to influence energy savings beyond their homes and daily jobs.

9 | From doing evaluation to evaluation use: importance of documenting data

The main conclusion from this part is:

KEY MESSAGE 10: good data is well-documented data.

This is illustrated by this quote from the case study on Energy Efficiency Agreements in Finland:

> "In reality, if two persons carry out impact evaluation of the same policy measure, they get different results. Even if I make the same calculation in successive years without proper documentation of the calculation method and definitions, the calculation can be different. This highlights the needs for good logic and documentation.
>
> Despite the need for good logic and documentation at the national level, based on personal experience on practicalities of evaluation I would not open the discussion for far reaching harmonization at the European level."

Documentation was also pointed in the case study on the municipal energy efficiency programme of the City of Vienna, as essential for reporting purposes:

> "The evaluation reports themselves are very detailed and are used as information sources for other programme evaluation reports (e.g. for the municipal Climate Protection Programme)."

This conclusion comes also from the difficulties encountered about clarifying the data found (scope, unit, etc.) when preparing the case studies. The documentation topic also raises the issue of harmonization. This was rarely discussed in the interviews of the case studies. However this issue was more frequently raised in the interviews done to assess stakeholders' needs and priorities, as well as in the discussions at the first EPATEE European workshop in Paris in October 2017.

Sections 9.1 and 9.2 present the different criteria used in the EPATEE case studies to document in a systematic way the data about energy savings and costs respectively. This shows the variety of metrics used to report energy savings, and the diversity in the scope of costs (when cost data could be found).

The documentation of energy savings has been particularly highlighted in cases of umbrella programmes (e.g., Municipal energy efficiency programme of the City of Vienna, German Energy Efficiency Fund) that include many distinct measures. In these cases, the use of common rules to document energy savings is essential to enable summing up the results from all measures to obtain the overall result of the umbrella programme.

The data themselves are gathered in summary tables in Annex I.

9.1 Criteria used to document data about energy savings

Table 24. Criteria used to document data about energy savings in the EPATEE case studies.

Case study	Final energy	Primary energy	First-year or annual	Cumulated annual	Lifetime cumulated	Energy unit	Normalized	Gross energy savings	Net or additional energy savings
[AT] Environmental Support scheme	X			X		MWh/y		X	
[AT] City Energy Efficiency Programmes of Vienna	X		X	X		GWh/y	X		X
[BE] Primes Energie	X		X			GWh/y	X	X	
[CR] Energy renovation of public sector buildings	X		X	X		PJ/y	X	X	
[CR] Individual heat metering in multifamily buildings	X		X	X		PJ/y	X	X	
[DK] EEO scheme	X		X			TJ/y	X		X
[FI] Energy Efficiency Agreement for Industries	X		X	X		GWh/y	X	X	
[FI] Voluntary energy audits for municipalities	X			X		GWh/y	X	X	
[FR] Voluntary agreement for freight companies	X		X	X		tCO_2/a		X	
[FR] "Future Investments" programme	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
[DE] Energy Efficiency Networks Initiative	X	X		X		PJ/y		X	
[DE] Energy Efficiency Fund	X	X		X		GWh/year	X	X	X
[IE] Better Energy Homes	X		X	X		GWh/year	X	X	
[IT] White Certificates Scheme		X	X	X		Toe/year	X		X
[LT] Renovation programmes with EU funding	X		X			GWh/year	X	X	
[NL] Multi-year agreements in the industry		X		X		PJ/year	X	X	
[NL] Subsidy scheme for housing corporations	X		X	X		million m3 gas/y	X	X	
[NL] Fiscal incentives for cars	X			X		PJ/year	X		X
[Nordic Countries] Nordsyn	X			X	X	GWh/y or GWh	X	X	
[UK] Supplier Obligations	X		X		X	TWh/y or TWh	X	X	
[UK] Warm Front	X			X		TWh/y	X	X	
[US] WAP	X		X			MMBtu/year	X	X	X

9.2 Criteria used to document data about costs/investments

The first table below provides an overview of the type of data about costs for which information could be found. Then the second table brings complementary details about the scope of these data.

Table 25. Criteria used to document data about costs/investments in the EPATEE case studies.

Case study	Public aids	Other public costs	Stakeholders' costs	Participants' costs	Total investments	Other cost data
[AT] Environmental Support scheme	F		n.a.			F
[AT] City Energy Efficiency Programmes of Vienna						
[BE] Primes Energie	F		n.a.		F	
[CR] Energy renovation of public sector buildings			n.a.		M	F
[CR] Individual heat metering in multifamily buildings	F		n.a.		F	
[DK] EEO scheme	n.a.	F	F			M
[FI] Energy Efficiency Agreement for Industries	n.a.	F	n.a.		F	M
[FI] Voluntary energy audits for municipalities	F	F	n.a.	F	M	M
[FR] Voluntary agreement for freight companies	n.a.	F	n.a.	F	F	
[FR] "Future Investments" programme	F		n.a.	F	F	
[DE] Energy Efficiency Networks Initiative	n.a.	M	n.a.	M		
[DE] Energy Efficiency Fund	F	M		M	M	
[IE] Better Energy Homes	F	F			F	
[IT] White Certificates Scheme	n.a.	F	F			
[LT] Renovation programmes with EU funding			n.a.		F	F
[NL] Multi-year agreements in the industry		F	n.a.			
[NL] Subsidy scheme for housing corporations	F		n.a.			
[NL] Fiscal incentives for cars	F		n.a.			
[Nordic Countries] Nordsyn		F	n.a.			F
[UK] Supplier Obligations	n.a.	F	F	M		M
[UK] Warm Front	F	M		M	M	M
[US] Auctions for capacity markets	F					
[US] Weatherization Assistance Program	F		F		F	

M: evidence that the data is monitored
F: data could be found and included in the case study
n.a.: not applicable

Table 26. Complementary information about the type of data on costs/investments.

Case study	Type of cost data
[AT] Environmental Support scheme	Amount of subsidies Environmentally relevant investment costs (marginal investments for performance beyond regulations)
[AT] City Energy Efficiency Programmes of Vienna	As the SEP is an umbrella programme including a variety of different schemes it is not possible to determine programme costs. This is also due to the fact that many of the instruments have their own source of funding.
[BE] Primes Energie	Public aids: budget commitments based on the applications received (paid amounts may be smaller due to withdrawals or non-compliance) Total investments: costs of the works supported (based on the invoices)
[CR] Energy renovation of public sector buildings	Total investments: total contracted value for the energy performance contracts (amount without VAT) Other costs: annual savings on energy bills for the renovated buildings
[CR] Individual heat metering in multifamily buildings	Public aids: amounts of grants awarded by the Croatian Energy Efficiency Fund Total investments: total investments made to install individual heat meters or heat cost allocators in apartment buildings
[DK] EEO scheme	Other public cost: administration costs for the Danish Energy Agency (management of the scheme + M&V) Stakeholders' cost: Costs reported by the energy distributors = incentive costs (energy advice, grants to final customers, subcontracting) and administration costs (quality control system, documentation, reporting) for all energy distributors (except heating oil distributors, as they don't recover their costs on the tariffs for energy networks) Other cost: investments by participants assessed on a sample of projects (in the ex-post evaluations) (total investment done by all participants is unknown)
[FI] Energy Efficiency Agreement for Industries	Other public cost: cost for administration, supporting participants in implementation, communication and marketing as well as monitoring and evaluation of the energy efficiency agreements (contract between the ministry and Motiva) Total investment: total investments made by the participants in industry and the private services (possibly including public subsidies received from other schemes) Other costs: annual cost savings from the actions implemented (based on data from the energy audits)
[FI] Voluntary energy audits for municipalities	Public aids: annual amounts of subsidies Other public cost: administration costs (not limited to audits in municipalities, but related to audits in all sectors + energy efficiency agreements), and costs of the web platform (initial investment + regular annual cost) Participants' cost: part of the cost of energy audit paid by the participants Total investment: investment costs of the actions recommended in the audits (based on the ex-ante estimates reported in the audits) Other cost: annual energy cost savings of the actions recommended in the audits (based on the ex-ante estimates reported in the audits)
[FR] Voluntary agreement for freight companies	External expenses (other public cost): all external expenses made by public bodies for the programme, including subcontracted studies, additional staff in regions, costs of events, … Internal expenses (other public cost): permanent staff involved in the programme (estimated in full-time equivalents by the different services, and assuming average full wage of 70,000 euros per full-time equivalent) Participants' cost: part of the support actions and human rsources not funded with public budget

Case study	Type of cost data
[FR] "Future Investments" programme	Public aids: average annual amount of public aid (subsidies, refundable aids and equity) Participants' cost: part of the investments paid by the participants Total investment: average annual provisional investments reported by project holders (note: investments are reported when applying for aids, but are usually done over several years) Data monitored by type of public aid, categories of project, and type of beneficiary organisations (distinguishing large companies, SMEs and others)
[DE] Energy Efficiency Networks Initiative	Other public cost: administrative costs of the networks (financed by the network participants) Participants' cost: network participation costs vary between 1000 and 5000 Euros per company and year. Each network has organised an average of four meetings per year for which network moderators invested about 20 working hours per meeting. In some regions or federal states, participant companies can apply for financial support in separate programmes.
[DE] Energy Efficiency Fund	Public aids: amount of money assigned to the grant measure "support of cross-cutting technologies" from the energy efficiency fund, in cumulated terms Other public cost: administrative cost calculated from average minutes spent for processing of each action in two action categories (simple and systemic) and weighting for cost of labour and overheads. Participants' cost: Part of the investments paid by the participants Total investment: Total investment of the projects receiving an aid from the Fund
[IE] Better Energy Homes	Public aids: annual amounts of grants paid Other public cost: administration costs (about 6% of the total public costs) Total investments: total material and labour costs for the actions receiving a grant
[IT] White Certificates Scheme	Other public cost: cost for the management and evaluation activities done by GSE paid since 2015 through a fee due for each request for white certificates Stakeholders' costs: costs recovered by the obligated gas and electricity distributors through tariff components. There is no official data on the global value of the investments done by ESCOs and end-users
[LT] Renovation programmes with EU funding	Total investment: compensation for soft loan interest exceeding 3%; the costs of investment plan, energy certification before and after renovation, technical project, supervision and expertise and administration; the costs of the renovation works (material and labor costs) and for additional heating system adjustment. Other cost data: average renovation cost/m² (calculated each year)
[NL] Multi-year agreements in the industry	Other public costs are the execution costs and refer partly to the costs for the NL Agency (i.e. the hours worked by NL Agency staff), partly to the costs for consultants that executed studies for the participants (scans, efficiency plans etc.) and for a small percentage direct subsidies to companies.
[NL] Subsidy scheme for housing corporations	Public aids: amounts of subsidies given to the housing corporations
[NL] Fiscal incentives for cars	Public aids: impact of the measure on State budget (about €1 billion less revenue from purchase tax on cars)
[Nordic Countries] Nordsyn	Other public costs: the budget for all the appliances tested between 2011 and 2013 (3 years) was estimated to be around 2.1 million Euro for all Nordic countries. Other costs include for example costs of documentation/reporting, but could not be estimated. Other costs: • average costs of 5.440 EUR per appliance tested (including administration, purchasing the products and performing the tests) • cost for the external study done to evaluate the scheme (about €27000)

Case study	Type of cost data
[UK] Supplier Obligations	Other public cost: Administration costs for Ofgem (management of the scheme + M&V): about 2m €/a for 2008-2012, now closer to 6m €/a Stakeholders' cost: Costs estimated/ reported by the energy suppliers = incentive costs (subsidies) and administration costs incurred by the energy suppliers (quality control system, documentation, reporting) (data for scheme year 2017-2018) Participants' costs: Investment by participants are assessed on a sample of projects (in the ex-post evaluations but high uncertainty around this data) (total investment by participants is unknown) Other costs: Costs per type of actions is estimated through surveys and stakeholder interviews (ex-post evaluations report high uncertainties around this data)
[UK] Warm Front	Public aids: Scheme expenditures = grants (83% of the expenditures in 2007/2008 (NAO, 2009)) + administration costs (fees of the scheme manager; 9%) + supporting services provided to the beneficiaries (surveys, post-installation inspections and benefit entitlement checks; 8%) Other public costs: funding brought by other public bodies than the government (e.g., local authorities) Participants' costs: part of the costs of the energy efficiency actions paid by the participants (from 75 to 97% of the actions (other than CFLs) could be fully funded by the grants, depending on the years) Total investment : full costs of the energy efficiency actions receiving a Warm Front grant Other costs: average cost per action type
[US] Auctions for capacity markets	Public aids: capacity payments are not officially reported, but estimated based on clearing price in different auctions and reported capacity cleared. These amounts include incentive costs (subsidies) and administration costs (verification, documentation, reporting).
[US] Weatherization Assistance Program	Public aids: total DOE funds invested in the weatherization of home without leverage Stakeholders' costs: funds from other sources (e.g., states, utilities) invested in the weatherization of home also receiving DOE funding Total investment: total expenditures on units weatherized that included DOE funding. From the Programme Year 2008 total funding, approximately 70% of the funds were spent on energy conservation measure installation, 10% on health and safety measures, 7% on audits and inspections, 12% on program management, and 1% on training and technical assistance. Use of funding is monitored by funding sources (DOE, state, utilities, etc.) and per Grantee and Subgrantee.

10 | From doing evaluation to evaluation use: importance of communication

The main conclusion from this part is:

KEY MESSAGE 11: don't neglect discussion and communication about evaluation results.

This point was not directly in the scope of the EPATEE case studies, but it arises from several of the interviews. So it was not systematically covered in all case studies. This will be further investigated in the next phase of the EPATEE project, in the task about how to integrate evaluation into the policy cycle.

First feedback collected in the case studies provides a good starting point to include the communication issue in this topic of integrating evaluation into the policy cycle. It shows that the discussion and communication about evaluation results can be as important as doing evaluation.

Table 27. Feedback collected about communication issues.

Case study	Details on communication about the evaluation (when mentioned in the interview)
[AT] Environmental Support scheme	Feedback from Michael Aumer (presentation at the EPATEE workshop in Vienna): It can sometimes be difficult to explain analysis and results in the evaluation report, and the format of the evaluation results and conclusions should be adapted to the audience targeted. To illustrate this point, Michael Aumer showed the thickness of the complete evaluation report, and then of the summary report that was distributed to the Members of the Parliament. Both formats enable discussions at expert and political level respectively. Discussions with the Parliament are very important, as the Parliament can decide changes on the details included in the Law establishing the programme, particularly who is eligible and the budget of the programme.
[AT] City Energy Efficiency Programmes of Vienna	The evaluators included one recommendation about communication: disseminating the results of the programme, both within the municipalities and towards the general public, to raise awareness.
[BE] Primes Energie	"The government changed between the time when the evaluation was commissioned, and the time when the evaluation results were released. This made that the evaluation results were taken as an assessment of the previous government, while they were based on the evidence available and did not include a political dimension. At the end, the new government chose to implement some of the recommendations made in the evaluation, depending on the priorities newly set."
[CR] Energy renovation of public sector buildings	"There is a question of indicators and which indicators will be the best factors for analysis in the evaluation process. (…) These indicators should not incorporate just the essential component, which is money spent per kilowatt hour (Croatian kuna or Euro per kWh), but also kWh per person living, residing or working full-time in the housing/public sector building. Further indicators in the long-term would be the impact assessment based on the calculation of a ratio of expenses on energy in an average household or public building. The number of educated, qualified work force per saved kilowatt hour would be another indicator to consider in order to make an evaluation successful and indicative of further steps to take in the energy efficiency framework. "
[CR] Individual heat metering in multifamily buildings	"As the representative of dwelling owners, I am monitoring these effects through data about the heat consumption and related bills. All tenants are informed about the effects through our building Facebook profile. This way, the awareness on energy consumption and consequences of our investment activities, but also our behaviour as energy consumers, is raised."

Case study	Details on communication about the evaluation (when mentioned in the interview)
[DK] EEO scheme	"It is always very motivating for us to know that the evaluation customer is genuinely interested in our results and that the results are used as basis for decisions regarding the policies that were evaluated. This also means that the stakeholders show a keen interest in the results. Our approach is when possible to engage the relevant stakeholders in the evaluation and to be open about the progress of the evaluation work so that the end result does not come as a shock. In my opinion, this approach benefits the overall decision-making process." "An additional advantage is that when we have had regular contact and discussion with the individual stakeholder groups, then all parties – us included – are better prepared for the media attention and can minimise the distortion of the evaluation results by the media. This is something which frequently happens."
[FI] Energy Efficiency Agreement for Industries	A very important indicator for the participants themselves is the cost savings achieved through participation. It is also a matter of public interest and getting attention in major media. "The success factors of this well-working policy measure have been good monitoring and evaluation, strong results and communication of results. This has led to increasing motivation and further improving results, i.e., a circle of positive development has been created. There is a wide positive consensus, all the way up to the ministries and ministers."
[FI] Voluntary energy audits for municipalities	"Comprehensive and reliable data was very important during the first years of the scheme for marketing purposes. If the audience is, e.g., managers of city hotels, then the results and experience presented must be from hotels located in city centres."
[FR] Voluntary agreement for freight companies	All evaluations supported by ADEME are reviewed by ADEME's evaluation committee in order to monitor the implementation of recommendations but operational decisions to ensure this implementation are not necessarily taken into account.
[DE] Energy Efficiency Networks Initiative	"Due to the large size of the networks initiative, reporting obligations to the EU in EED Article 7 are in place. Furthermore, in the political arena, critical inquiries arrive from the press, societal organisations and the general public."
[DE] Energy Efficiency Fund	"Sometimes, certain sentences need to be revised in their formulation or their content. In general though, discussions between the evaluators and the ministry are open and based on mutual understanding." "The big question is always how to use scientific results in practice?"
[IE] Better Energy Homes	"The ex-post evaluation was very well accepted by the Ministry. Indeed the evidence brought by the evaluation changed their perception of the scheme. There were no more questions about the rationale or interest to implement this scheme. At the opposite, the questions were about how to make the scheme grow. So evaluation needs move from justifying the scheme to understanding how to tackle the difficulties to get a higher participation. SEAI has indeed now a team specialised in behavioural sciences. Their work aims at improving the communication about the scheme, to use the right message for the right target." The ex-post evaluation was also used to prepare a public communication showing to households the actual energy savings achieved thanks to the programme.
[NL] Multi-year agreements in the industry	The NL Agency uses the data reported by the companies to inform the Energy Efficiency Consultative Group (OGE) and the Dutch House of Representatives about the progress of the scheme.
[Nordic Countries] Nordsyn	Results of the evaluation confirmed the cost-effectiveness of Nordsyn actions (market surveillance cooperation), demonstrating with a cost-benefit analysis the impact of implementing market surveillance. The results of the evaluation were communicated to the Nordic Council of Ministers that thus adopted a budget to continue Nordsyn over 2016-2017.
[UK] Warm Front	"The research project delivered very rich materials, in various fields. About 20 academic papers were published. And a synthesis report summarized the main results" "In general, politicians do not want to be told that their policy is a failure. Their focus, in terms of evaluation, is to know the results of the policies in order to communicate about them. This makes that less attention is paid to the question of why and how these results were achieved (or not)."

Case study	Details on communication about the evaluation (when mentioned in the interview)
[US] Weatherization Assistance Program	"Other stakeholders are interested as well in the WAP evaluations. For example, the US Congress uses them to see if this is a good program, worth funding. NASCSP (National Association of State Community Services Programs) uses them to see how well the program is working, and as a basis for discussions with its members about possible improvements." "The WAP community is broader than the people working directly on WAP. It is important to keep the whole community informed about ongoing evaluation activity and then about evaluation results. Numbers alone don't tell the full story. Part of the evaluation was about process evaluation, bringing qualitative aspects that are essential to explain what happened, make sense of the numbers and put them in the right context." "One "difficulty" for the evaluation team was that the evaluation findings did not support the savings values that were being reported by DOE. Even though the evaluated estimates of savings documented that the WAP program was a "high performing" energy efficiency program when compared to other low-income and market rate energy efficiency programs. The actual savings simply did not match DOE's reported values. After the evaluation team presented the preliminary "disappointing" results to ORNL and DOE, DOE made the decision that the evaluation team was not allow to present the study findings on energy savings to the broader research community until all of the study reports were complete. That made it difficult for the study authors to gain insights on the findings and potentially improve the reports by hearing from other researchers about how the study findings compared to their own results." "In our work with state grantees, we often communicate the findings of the evaluation and make recommendations to those state grantees on how to design and implement their programs."

11 | Other lessons learnt and points of debates raised in the case studies

In addition to the issues systematically covered in all EPATEE case studies, other and more specific issues might have been raised, as reported in the table below.

Table 28. Other lessons learnt identified in the EPATEE case studies.

Case study	Other lessons learnt
[AT] Environmental Support scheme	Importance of the preparatory work done by the evaluation commissioner (to prepare information to be provided to the evaluators) Difficulties to find evaluators meeting the criteria (independence, knowledge of the scheme, technical and economic expertise) Need to adapt the evaluation objectives to the budget available for the evaluation
[AT] City Energy Efficiency Programmes of Vienna	The evaluation of some measures (e.g. soft measures) was not possible due to missing or too complex methodologies. Thus standardised calculation methods (e.g. according to ESD an EED) were helpful to quantify the energy savings and to limit the effort for evaluation.
[DK] EEO scheme	Difficulties to get market data, or assess impacts on markets of EE actions "In cases where, for example, the stakeholders have strong opposing interests, it can be useful to give presentations for each group separately so that there is room for free discussion. This also provides us as evaluators with important insights regarding the policies and how they actually operate, which in turn benefits the quality of the evaluation results. Our experience is that the stakeholders are very positive towards such an approach." "With regard to the role of evaluators, one "guideline" that can be illustrated by the 2008 evaluation is that evaluators should remember to take a critical look at the evaluation scope defined in the tender. In 2008, we suggested that the scope should be expanded to also include the energy taxation scheme. Doing so enabled us to clearly compare the impact and cost-effectiveness of each of all the energy efficiency policies."
[FR] "Future Investments" programme	Lessons learnt about the difficulties to perform econometric analyses on a programme supporting R&D projects
[DE] Energy Efficiency Networks Initiative	The participants have an interest to know their data to be secure and not disclosed to other companies, nor a federal ministry of other public offices. This requires a special attention with data collection procedures, storage and access.
[DE] Energy Efficiency Fund	Energy data are given as both final energy and primary energy data. In the light of rapidly changing means of energy production, final energy falls short of showing these changes. The German Ministry of Economic Affairs is therefore mostly interested in primary energy data. Importance of defining a common methodological, monitoring and reporting framework to ensure consistency in the distinct evaluations of the measures included in the Fund.
[IT] White Certificates Scheme	The challenge due the growing share of projects in industry since 2010. These projects are often very specific. This implies a higher degree of difficulty in dealing with M&V protocols, consumption baselines, and additionality (compared to projects in the residential or service sectors). Proponents are thus requests to report a large amount of valuable information. Unfortunately, the original structure of the database is not adequate to such detailed information and thus the evaluation activities are complex. It is advisable for new schemes to pay the due attention to this aspect to facilitate indexing and analysis activities.
[NL] Fiscal incentives for cars	The cost of the evaluation done by PBL (ca. €100,000), as percentage of the total costs of the greening of the purchase tax system, is negligible.
[UK] Supplier Obligations	"It is difficult to say whether we should do an evaluation at a specific point as scheme keeps changing. Embedding an evaluation process that is continuous can accommodate for that. We also base our evaluation work on the so-called Magenta Book which contains broad guidelines for evaluation to be applied across all departments in the UK."

Case study	Other lessons learnt
[UK] Warm Front	"One difficulty we encountered in the research project was that the policy changed over the duration of our studies, partly due to our results. This meant that we were not totally in the conditions for an academic methodology where researchers are supposed not to be affecting the object they are investigating. Using an action-based methodology means assuming that the researchers can have an influence on the policy and/or on the results they investigate. Researchers get involved in the process, and thereby have more interactions with stakeholders and participants. Which makes possible to identify problems sooner. Using large samples and long-time series is not the most appropriate way in all cases. For example, there is no need to use large sample to find that the interventions are not done in the right order, creating risks on airtightness. This can be found with a limited number of interviews, and this can be directly used to improve the scheme." "Our research team has been championing the approach of "energy epidemiology". The idea is to take back the concept of "epidemiology" from the health sciences. Initially, epidemiology stood for "study of a population and what happens in a population". Extensive use of epidemiology by health sciences made that the concept is now most often understood as study of a population to investigate health impacts in a population. The approach of energy epidemiology focuses on empirical research, using large population datasets. This research can for example be used to analyse the differences between energy consumption as measured in laboratory tests and energy consumption as measured in field conditions."
[US] Auctions for capacity markets	In the ISO-NE Forward Capacity Market, EE capacity is eligible to participate as long as it is operational. Therefore, capacity offer of existing EE resources needs to be updated in the application stage for each annual forward auction to account for changes in technical performance or operational practice. "Forecasting future capacity needs is more complex when energy efficiency can participate as a resource because the need for capacity is based on load forecasts, which will be impacted by successful energy efficiency programs. These impacts must be accounted for correctly in load forecasts to avoid under- or over-procuring capacity in future periods."

Table 29. Other lessons learnt and points of debates raised in the EPATEE case studies.

Case study	Points of debates
[AT] Environmental Support scheme	External evaluation vs. knowledge of the scheme
[BE] Primes Energie	How to assess additionality/net savings Internal vs. external evaluation
[CR] Energy renovation of public sector buildings	Scaled savings vs. metered savings
[CR] Individual heat metering in multifamily buildings	Cost-effectiveness of the actions (heat cost allocators) in view of revising the regulation
[DK] EEO scheme	How to assess additionality/net savings (and reliability of results from surveys)
[FI] Energy Efficiency Agreement for Industries	Internal vs. external evaluation Harmonisation
[DE] Energy Efficiency Networks Initiative	Balance between making participation as easy as possible (lower requirements) and reliability of the results (higher monitoring requirements)
[DE] Energy Efficiency Fund	Use of surveys to assess free-rider effects.
[IT] White Certificates Scheme	Finding the good balance between the level of requirements (monitoring, additionality) and the resulting costs.

Case study	Points of debates
[NL] Multi-year agreements in the industry	Bottom-up vs. top-down approach to monitor the result of the scheme.
[NL] Subsidy scheme for housing corporations	What methods to use when assessing actual energy savings and comparing with deemed savings
[NL] Fiscal incentives for cars	Reliability and relevance of values from laboratory test vs. field measurements
[US] Weatherization Assistance Program	Debate about evaluation results (as several evaluation studies were done in parallel of the official ones) Debate about independency: “OMB, the US Office of Management and Budget, was correct in suggesting that ORNL would have an apparent conflict of interest in conducting the WAP evaluation. In their work on the Health Benefits of WAP, ORNL rejected use of the standard "difference in difference" analysis approach. Rather, they implemented an alternative approach that attributed much higher health benefits to the WAP program, without furnishing a detailed explanation of why they were rejecting the standard approach. As such, they give the appearance of having a conflict of interest in their research.”

Annex I: Overview of data collected in the EPATEE case studies

Data about targets

Country	Name of the measure	starting date	ending date	Expected energy savings / results	benchmark
Austria	Aid for environmental protection measures (UFI)	1986	on-going	25.1 PJ of final energy savings cumulated over 2014-2020 from actions implemented in 2014 and 2015	15% of the target for EED art.7
Austria	Vienna Energy Efficiency Programme (SEP)	2006	2015	rate of new annual final energy savings (or first-year savings) of 180 GWh/y on average between 2006 and 2015	Total final energy consumption in Vienna amounted to around 38,000 GWh in the last years on average
Belgium (Wallonia)	Primes Energie (grants for energy renovation)	2004	on-going	2.07 TWh/y (7.4 PJ/y) in 2020 from actions over 2014-2020 (cumulated annual final energy savings) Average expected rate of new annual savings = 0.3 TWh/y	32% of the EED article 7 target (for 2014-2020) Expected rate of new annual savings = 1.3% of 2013 space heating consumption for households in Wallonia
Croatia	Energy renovation programme for public sector buildings	2014	2018	final annual energy savings of about 0.23 PJ/year in 2020 from actions implemented in 2014-2015	total energy consumption of the public sector buildings: 1.63 PJ/y
Croatia	Individual heat metering in multi-family buildings	2014	2016	new final annual energy savings of about 0.4 PJ/y each year over 2014-2016	Assumption that in approximately half of the dwellings connected to a central heating system (total number of such dwellings is approx. 155,000), i.e. in about 75,000 dwellings, implementation of individual heat allocators is feasible in period 2014-2016. That amounts to 25,000 dwellings annually.
Denmark	EEO scheme	2006	on-going	83.9 PJ/year (23.3 TWh/year) in 2020 from actions over 2014-2020 (cumulated annual final energy savings)	100% of the target for EED article 7 annual target for 2016 = 2.6% of Danish 2014 final energy consumption (excluding transport)
Finland	Energy Efficiency Agreement for Industries	1997	on-going	For 2020: 200 GWh/a in the private services sector, 770 GWh/a in mid-sized industry and 11 691 GWh/a in energy-intensive industry (cumulative annual final energy savings from actions implemented from 1997 and still operating in 2020)	64% of the target for EED art.7; Cumulative annual energy savings in industry estimated for 2016 (from actions over 1997-2016 and still operating in 2016) account for 8.0% of the sectoral total final consumption in 2016

Country	Name of the measure	starting date	ending date	Expected energy savings / results	benchmark
Finland	Energy audits in municipalities	1992	on-going	Annual final energy savings expected in 2020: 88 GWh/a	This measure is estimated to account for 20% of all energy savings in the municipal sector in 2020. However, the contribution to the EED article 7 target in 2020 is very small.
France	"Investments for the Future" programme	2010	2020	No target in terms of energy savings (measure supporting R&D&I, with the objective to foster the emergence of innovative products/services)	n.a.
France	Voluntary agreement for freight companies	2008	on-going	no fixed target in terms of energy savings for the programme but companies committing to the programme must set themselves objectives over a 3-year period	n.a.
Germany	Energy Efficiency Fund (data for the sub-measure "support for highly efficient cross-cutting technologies in SMEs")	2011	on-going	2,260 TJ (628 GWh) per year in 2020 from actions implemented over 2012-2020 (cumulated annual final energy savings)	0.46% of German 2020 reduction goal
Germany	Energy Efficiency Networks Initiative	2014	2020	75 PJ (20,8 TWh) Primary energy and 5 Mt CO2-eq.	
Ireland	Better Energy Homes	2008	on-going	1324 GWh/year of final energy savings in 2020 for actions carried out between 2011 and 2020	target = 18% of the annual final energy savings expected in 2020 from buildings
Italy	White Certificates Scheme	2005	on-going	4.3 Mtoe/year (final energy) in 2020 from actions implemented over 2014-2020	60% of the national target for EED art. 7
Lithuania	Renovation programme with EU funding	2005	on-going	About 1 TWh/y of final annual energy savings in 2020 for actions implemented over 2005-2020	The target represents the renovation of 4000 apartment blocks (120.000 dwelling units)
Netherlands	Subsidy scheme for housing corporations in Amsterdam	2011	2014	About 3 Mm3 of gas saved/year from actions implemented over 2011-2014 (16,500 label steps	The dwelling stock of the Amsterdam housing corporations includes 160,000 dwellings.

Country	Name of the measure	starting date	ending date	Expected energy savings / results	benchmark
Netherlands	Multi-year agreements in the industry	2001	on-going	To support the businesses that signed the agreement to reach a 30% improvement in energy efficiency from 2005-2020.	The 946 companies that reported their results for LTA3 (non-ETS) in 2016 consumed 247 PJ (primary energy) in 2016. They represented about 23% of the total energy consumption in industry, and about 7% of the total primary energy consumption in the Netherlands
Netherlands	Purchase tax on passenger cars	2006	on-going	Expected cumulative final energy savings over 2014-2020 for EED art. 7 from all measures to promote efficient cars were estimated to range between 16 and 28 PJ	Dutch target for EED art.7: 482 PJ of cumulative final energy savings over 2014-2020 (the expected results from the measures for efficient cars represented about 3 to 6% of the Dutch target). Share of energy consumption of passenger cars in in the final energy consumption of Netherlands: about 12 % (250 PJ/2000PJ).
Nordic Countries	Nordsyn (market surveillance cooperation)	2013	2017	Nordsyn does not include directly quantitative objectives in terms of energy savings. For example, expected energy savings from the implementation of ecodesign requirements for products were estimated for Denmark to 5,640 GWh of final energy per year in 2020	The expected energy savings in 2020 from the implementation of the Ecodesign requirements in Denmark corresponded to 5 % of Danish final energy consumption in 2011 excluding transport.
UK	Supplier Obligations	1994	on-going	4.4 TWh/y (15.8 PJ/y) in 2020 from actions over 2014-2020 (cumulated annual final energy savings)	7% of the target for EED article 7 annual target for 2016 = 0.2% of UK 2016 residential final energy consumption
UK	Warm Front	2000	2013	8.4 TWh/year in 2020 from actions over 2000-2013	About 11% of English households were assisted by the scheme (over 13 implementation years).
US	Auctions for capacity markets	2010	on-going	17 GW in 2020 from actions over 2010-2020 (cumulated capacity cleared in the auctions)	Energy efficiency received more than 6% of all capacity payments awarded in the 2017 auction
US	Weatherization Assistance Program	1976	on-going	No official quantitative target set in terms of expected energy savings. The objectives are usually expressed in terms of number of dwellings to be weatherized per year.	WAP's annual appropriation has been supporting the weatherization of approximately 100,000 homes. Approximately 35 million households were eligible for WAP in PY (Programme Year) 2008 (i.e. about 30% of all U.S. households).

Data about means and outputs

Country	Name of the measure	annual public budget (M€/y)	annual stakeholders' cost (M€/y)	annual investments (M€/y)	Explanations
Austria	Aid for environmental protection measures (UFI)	75		623	Public budget: annual amount of subsidies averaged over 2011-2013 Annual investment: annual amount of marginal investments (extra cost beyond regulations) averaged over 2011-2013
Austria	Vienna Energy Efficiency Programme (SEP)				As the SEP is an umbrella programme including a variety of different schemes it is not possible to determine programme costs. This is also due to the fact that many of the instruments have their own source of funding.
Belgium (Wallonia)	Primes Energie (grants for energy renovation)	37		125	Public budget: amount of grants in 2013 Annual investments: annual amounts averaged over 2008-2012
Croatia	Energy renovation programme for public sector buildings				Total contracted value (signed over 2014-2016): €99.4 million, for 21 Energy Performance Contracts (68 buildings; 225,000 m²)
Croatia	Individual heat metering in multi-family buildings	2		5.3	Annual investments: amount of total investments averaged over 2014-2016 Public budget: amounts of grants averaged over 2014-2016
Denmark	EEO scheme	0.54	210		Data for 2015 (public budget = administration costs ; stakholders' cost = cost for the obligated parties)
Finland	Energy Efficiency Agreement for Industries	1		66	Public budget: data for 2017, operational costs of the agreement scheme's administration Annual investments: data for 2016 (investments for actions in industry + private services)
Finland	Energy audits in municipalities	0.43	0.43		Public budget: average annual amount of subsidies over 2008-2016 (ranging from 0.13 to 0.69 M€/y) Stakeholders' cost: subsidy rate of 50%, therefore 50% of the audit costs to be paid by the participants

Country	Name of the measure	annual public budget (M€/y)	annual stakeholders' cost (M€/y)	annual investments (M€/y)	Explanations
France	"Investments for the Future" programme	300		900	Public budget: average annual amount of public aid (subsidies, refundable aids and equity) over 2010-2017 Annual investment: average annual provisional investments reported by project holders (note: investments are reported when applying for aids, but are usually done over several years)
France	Voluntary agreement for freight companies	0.82	n.a.	1.02	Public budget: external and internal resources averaged over the 2008-2012 period Annual investments: total costs (including staff) for public and private bodies, averaged over the 2008-2012 period
Germany	Energy Efficiency Fund (data for the sub-measure "support for highly efficient cross-cutting technologies in SMEs")	20			Public budget: annual amount of grants approved for the measure "support for highly efficient cross-cutting technologies in SMEs", averaged over 2012-2017 (variation from 10 M€ in 2012 to 84 M€ in 2015)
Germany	Energy Efficiency Networks Initiative				Participants' cost: network participation costs vary between 1000 and 5000 Euros per company and year.
Ireland	Better Energy Homes	17	energy companies and local authorities may also help	69,5	Public budget: amount of grants (data for 2016) Annual investments: total costs of the actions (data for 2016)
Italy	White Certificates Scheme	6,5	700		Public budget: cost of management and monitoring activities of GSE in 2016 Stakeholders' cost: cost recovered by electricity and gas distributors through regulated tariff components in 2016
Lithuania	Renovation programme with EU funding			41	Annual average based on total investment over 2005-2013
Netherlands	Multi-year agreements in the industry	5			Public budget: costs for the NL Agency (i.e. the hours worked by NL Agency staff), part costs for consultants that executed studies for the participants (scans, efficiency plans etc.) and for a small percentage direct subsidies to companies (estimation of annual costs in 2017). There can be significant differences between the years.
Netherlands	Subsidy scheme for housing corporations in Amsterdam	13,4			Public budget: subsidies given to the housing corporations (annual average over July 2011 – December 2013)

Country	Name of the measure	annual public budget (M€/y)	annual stakeholders' cost (M€/y)	annual investments (M€/y)	Explanations
Netherlands	Purchase tax on passenger cars	1000			Public budget: impact on the public budget (i.e. decrease in the revenues from the purchase tax on cars), estimate of the annual average
Nordic Countries	Nordsyn (market surveillance cooperation)	0.7			Public budget: budget for testing samples of appliances (estimated annual average for 2011-2013). Other costs include for example costs of documentation/reporting, but could not be estimated.
UK	Supplier Obligations	6	713		Public budget: Administration costs for Ofgem (management of the scheme + M&V): about 2m €/a for 2008-2012, now closer to 6m €/a Stakeholders' costs: Costs estimated/ reported by the energy suppliers = incentive costs (subsidies) and administration costs incurred by the energy suppliers (quality control system, documentation, reporting) (data for scheme year 2017-2018)
UK	Warm Front	288			Public budget: Scheme expenditures = grants (83% of the expenditures in 2007/2008 (NAO, 2009)) + administration costs (fees of the scheme manager; 9%) + supporting services provided to the beneficiaries (surveys, post-installation inspections and benefit entitlement checks; 8%) ; annual average calculated over the whole life of the scheme (13 implementation years), to be taken with caution as there were large differences from one period to another.
US	Auctions for capacity markets	170			Public budget: capacity payments are not officially reported, but estimated based on clearing price in different auctions and reported capacity cleared. These amounts include incentive costs (subsidies) and administration costs (verification, documentation, reporting). (annual average of estimated payment to EE resources over 2017-2020)
US	Weatherization Assistance Program	210	220		Public budget: total DOE funds invested in WAP for Programme Year (PY) 2008. Usually (before the recovery plan/ARRA period), annual DOE funding varied between €170 to 230 million/y. During the ARRA period, funding was drastically increased to more than €1700 in PY 2010. Stakeholders' costs: other funding for WAP from other sources (e.g., states, utilities) for PY 2008 (it was about €280 million for PY 2010).

Data about energy savings

Country	Name of the measure	reported energy savings	gross or net ?	Level 1 method	Level 2 method	Type of baseline	Adjustments, correction and other factors taken into account
Austria	Aid for environmental protection measures (UFI)	1.2 TWh/y from actions implemented over 2011-2013	gross	scaled savings	Method 5	“actual before”	No correction factors used Additionality criteria: performance > regulations; and payback time > 3 years (projects are additional, but savings are “gross” savings, as baseline = “actual before”)
Austria	Vienna Energy Efficiency Programme (SEP)	new annual savings of around 150 GWh/y on average over 2006-2014	additional	deemed and scaled savings	Method 3 and method 5	Mainly “stock average” or “market average”. And minimum energy performance standards for actions in existing buildings.	SEP savings are additional compared to minimum energy performance requirements enforced by EU and Austrian regulations. Double counting with Federal measures is avoided.
Belgium (Wallonia)	Primes Energie (grants for energy renovation)	179 GWh/y in 2013 for actions implemented in 2013	gross	scaled savings	method 5	"stock average" for the characteristics of the building components (updated regularly)	use of normalised weather conditions and behaviours; no other adjustment factor applied; performance criteria on actions to ensure performance additionality
Croatia	Energy renovation programme for public sector buildings	Final annual energy savings in 2016: 0.177 PJ/y from actions implemented from 2014 to 2016; New annual final energy savings for actions installed in 2016: 0.053 PJ/y	gross	scaled savings	method 5	“actual before”	normalization of weather conditions, occupancy rates and operating hours

Country	Name of the measure	reported energy savings	gross or net ?	Level 1 method	Level 2 method	Type of baseline	Adjustments, correction and other factors taken into account
Croatia	Individual heat metering in multi-family buildings	Final annual energy savings of 0.119 PJ/y in 2016 from actions implemented from 2014 to 2016	gross	Deemed savings	Method 3	“actual before”(energy consumption before the installation of the heat cost allocators)	normalization of weather conditions no assessment of rebound effect, but it is noted that most buildings were over-heated before the installation of heat allocators
Denmark	EEO scheme	10961 TJ/y (about 3 TWh/y) for first-year final energy savings achieved from actions implemented in 2016	net	deemed or scaled savings	method 3 or method 5	"before" energy consumption (except for replacement of equipment where repair work cost > 25% of replacement cost, then baseline = market average or legal requirement)	Deemed savings are normalized (e.g., weather conditions, heating behaviours); Scaled savings are adjusted for changes in operation hours, production volumes, etc.; Conversion factors (for substitution between energy sources); Reduction factors (based on additionality assessments done in previous ex-post evaluations); Prioritisation factors (to favour some action types, e.g. actions with longer lifetime).
Finland	Energy Efficiency Agreement for Industries	11.1 TWh/y achieved in 2016 from actions implemented over 2008-2016 and still operating in 2016	gross	scaled savings	method 5	"before" energy consumption (or minimum energy performance standards when actions covered by EcoDesign)	Double counting with other policy measures is tracked;
Finland	Energy audits in municipalities	89 GWh/y achieved in 2016 from actions implemented over 1995-2016 and still operating in 2016	gross	scaled savings	Method 5	"before" energy consumption (or “actual before” when consumption has been metered)	
France	“Investments for the Future” programme	Not available	net	Scaled savings	Method 5	Reference scenario	Additionality assessed when defining the reference scenario

Country	Name of the measure	reported energy savings	gross or net ?	Level 1 method	Level 2 method	Type of baseline	Adjustments, correction and other factors taken into account
France	Voluntary agreement for freight companies	1.6 million tons of CO_2eq avoided over 2008-2016	Gross	Top-down savings	Method 8	"before" energy consumption	Calculation based on energy consumption indicator per ton.km or per passenger.km (so normalisation for distances travelled and weight or passengers carried)
Germany	Energy Efficiency Fund (data for the sub-measure "support for highly efficient cross-cutting technologies in SMEs")	525 GWh/year in 2016 for actions implemented over 2012-2016 (cumulated annual final energy savings)	Gross (but net results also evaluated)	deemed or scaled savings	method 3 or method 5	"before" energy consumption	Free-rider effects determined based on ex-post surveys. Double counting (interaction effects between the different sub-measures of the Fund)
Germany	Energy Efficiency Networks Initiative	18.8 PJ/y (cumulated annual primary energy savings targets from networks launched over 2015-2017)	Gross	Metered, deemed or scaled savings	Methods 1, 2, 3, 4 or 5	"before" energy consumption, minimum standards or market average	The baseline is either minimum standards or before energy consumption or, whichever is higher. In case of a new installation a minimum standard is used. When not available market averages are used. No adjustment factors are applied. Additionality in the sense of the EED is addressed using the baseline of minimum standards to avoid double counting with existing minimum requirement policies.
Ireland	Better Energy Homes	Cumulative annual final energy savings: 994 GWh/year in 2016 (for actions implemented over 2009-2016)	gross	deemed savings	method 4	stock average (standard energy consumption per dwelling type)	Rebound effect (conservative values per type of dwelling, based on the comparison between modelled and metered energy consumption) Use of normalised weather conditions

Country	Name of the measure	reported energy savings	gross or net ?	Level 1 method	Level 2 method	Type of baseline	Adjustments, correction and other factors taken into account
Italy	White Certificates Scheme	about 2 Mtoe/y in 2016 (annual primary energy savings)	additional	deemed, scaled or metered savings	methods 1, 3 or 4	highest energy performance from legal requirements, market average and before situation	Adjustments for industrial production, weather, plant or building usage, etc. (for scaled and metered savings). Baseline defined (and verified) to ensure savings are additional. Double counting (verifying certificates are not issued twice for the same action)
Lithuania	Renovation programme with EU funding	About 200 GWh/y (new final annual energy savings) from actions implemented in 2016	gross	scaled savings	method 5	actual before	Use of standardized heating behaviours and weather conditions; No adjustment (rebound effect, free-rider effect, etc.) is applied
Netherlands	Multi-year agreements in the industry	63.4 PJ saved in 2016 vs. 2005 (primary energy savings, excluding actions on renewable energy)	gross	Deemed savings	methods 3 and 4	actual before	Reference year is 2005. Energy savings in the annual reports are gross energy savings: they do not include any causality assessment (i.e. possible free-rider or spill-over effects are not taken into account). In parallel, a decomposition analysis is done to show, among others, the changes in energy consumption due to factors other than the energy savings project.
Netherlands	Subsidy scheme for housing corporations in Amsterdam	About 0.9 Mm3 of gas saved/year from actions implemented over 2011-2014	gross	other	method 6	stock average	prebound effect (cases where, before implementing an energy efficiency action, end-users tend to consume less energy than estimated by engineering models) net impacts could not be evaluated
Netherlands	Purchase tax on passenger cars	11 PJ/year (final energy) from new cars sold over 2007-2015	net	Deemed savings	Method 3	Control group	rough estimate based the monitoring of the decrease in the specific emissions of new cars (in gCO_2/km) compared to the average trend in EU countries, and assuming a 50% rebound effect (the overall estimate was cross-checked with PBL evaluators)

Country	Name of the measure	reported energy savings	gross or net ?	Level 1 method	Level 2 method	Type of baseline	Adjustments, correction and other factors taken into account
Nordic Countries	Nordsyn (market surveillance cooperation)	over-consumption of about 18 GWh/year (or 168 GWh over products' lifetime) could be avoided in Nordic countries for sales of a typical year	lost energy savings	deemed savings (mix of ex-ante and ex-post)	method 4	minimum energy performance requirements	average annualized results from data over 2011-2013 It should be noted that 13 out of the estimated 18 GWh/year come from a single product category (combined fridge-freezers). Free-rider effects do not apply to Nordsyn, as it is about the implementation of a regulation. Spill-over and rebound effects were not taken into account in the evaluation.
UK	Supplier Obligations	about 10 TWh of final energy savings cumulated over the lifetime of actions implemented in 2015		deemed savings	method 3		pre-defined carbon/ energy savings ratios according to standardised types of actions taking into account in-use factors which account for rebound effects and performance gaps
UK	Warm Front	(1) cumulated annual final energy savings of 8.0 TWh/year, from actions installed over 2000-2010. (2) negligible changes in energy consumption due to high comfort taking	gross	scaled savings	method 5	"actual before"	(1) This was estimated based on the data from the annual monitoring. It corresponds to "theoretical" energy savings (result reported in NEEAP 2011). (2) This statement of NEEAP 2014 is consistent with the results of the qualitative survey of beneficiaries in the final process evaluation. However the results from this survey cannot be considered representative (small sample used to get qualitative insights).

Country	Name of the measure	reported energy savings	gross or net ?	Level 1 method	Level 2 method	Type of baseline	Adjustments, correction and other factors taken into account
US	Auctions for capacity markets			Metered savings	Method 1	"actual before" or "minimum performance standards"	Deemed (method 4) and scaled (method 5) savings can be used if complemented with metered data. The main indicators monitored are the capacity cleared and the verified capacity performance. Energy savings of the projects included in the bids are not evaluated by the capacity market scheme. However, as most of EE resources in the forward capacity market are part of a regulated utility EE obligation, their energy savings are evaluated under the regulatory framework of utility EE obligation instead.
US	Weatherization Assistance Program	2.4 PJ/y from actions of PY 2008 8.0 PJ/y from actions of PY 2010	net	Metered savings	Method 2	"Control group"	weather-normalized net energy savings (estimated by a statistical comparison of a sample of participants with a control group) in addition, an Indoor Environmental Quality (IEQ) study about works completed during the ARRA period concluded that there was no rebound effect relevant to home heating

Information about other indicators or aspects monitored or evaluated

Country	Name of the measure	Other impacts monitored and/or evaluated	Indicators about (cost-)effectiveness or efficiency of the policy measure	Other indicators/aspects monitored and/or evaluated
Austria	Aid for environmental protection measures (UFI)	Reductions in CO_2 emissions; Investments triggered; Results per region; Macro-economic impacts; Employment effects.	Public efficiency indicator (euro/tCO_2 avoided)	average processing time (time between application and decision on eligibility)

Country	Name of the measure	Other impacts monitored and/or evaluated	Indicators about (cost-)effectiveness or efficiency of the policy measure	Other indicators/aspects monitored and/or evaluated
Austria	Vienna Energy Efficiency Programme (SEP)			No quantitative indicators other that energy savings were included in the evaluation reports. But the evaluators noted that the next programme, SEP 2030, had to be compatible with all aspects of Viennese energy policy, whose key points are energy efficiency, environmental and climate protection, security of supply, economic efficiency and social justice. They also recommended to consider the interactions between energy efficiency and other policy objectives, mentioning for example affordable housing and urban planning. Moreover, the top-down analysis looked at the influence of economic and population growth and other factors on total final energy consumption
Belgium (Wallonia)	Primes Energie (grants for energy renovation)	Reductions in CO_2 emissions; Employment effects	Effectiveness assessed in terms of goal achievements, and public cost per energy saved (c€/kWh saved) (but these data are not included in the reports available)	distributive effects (differences in participation and types of actions depending on the income class); review of the scheme management (by the Court of Auditors); chronological analysis of the participation (to analyse how households react to changes in the scheme); needs in training for energy efficiency and building professionals (based on a survey of companies)
Croatia	Energy renovation programme for public sector buildings	Reductions in CO_2 emissions; Total investments; Employment effect; Geographical dispersion	Cost effectiveness of implemented actions are compared with cost optimal actions that are obtained by calculating costs and benefits for several predefined scenarios for each building	Analysis of financial challenges and organizational issues
Croatia	Individual heat metering in multi-family buildings	Reductions in CO_2 emissions;	Net present value of investment and energy savings	Analysis of complaints from end-users, and reasons in case of higher heating bills (after the installation of heat allocators)
Denmark	EEO scheme	Reductions in CO_2 emissions (however this result is not included in the evaluation reports)	Cost-effectiveness of the obligated parties (costs incurred by the obligated parties per reported kWh saved); Socio-economic cost-effectiveness (socio-economic net value of additional energy savings projects over their lifetime);	Costs per type of actions (monitored to assess trends related to costs for obligated parties); Qualitative assessment of the impact of the scheme on the energy efficiency markets (survey of contractors); Attempt to assess spill-over effects with an econometric analysis; Impact of the scheme on energy prices.

Country	Name of the measure	Other impacts monitored and/or evaluated	Indicators about (cost-)effectiveness or efficiency of the policy measure	Other indicators/aspects monitored and/or evaluated
Finland	Energy Efficiency Agreement for Industries	Reductions in CO_2 emissions; Use of renewable energy sources	Direct payback time, cost saving (on energy bills) and investment cost are monitored (based on data reported by the participants)	Public budgets used for energy audits and subsidies; Actions done using the ESCO model; Various indicators about the approach used by the participants (energy management, etc.); Satisfaction and feedback from the participants
Finland	Energy audits in municipalities	Reductions in CO_2 emissions; Water savings related to water heating	Direct payback time, cost saving (on energy bills) and investment cost are monitored (based on data reported in the energy audits)	Public budgets used for energy audits and subsidies; Audited building volumes by building types; Number of actions recommended and of recommendations implemented;
France	"Investments for the Future" programme	Reductions in GHG emissions; Job impacts; Leverage effect	At project level: Net Present Value. At programme level: Leverage effect, fiscal multiplier of public expenditure (BFTB: "Bang for the Buck")	Investments in R&D and in environmental fields + indirect impacts on participants' business development (production, turnover, employment, etc.). Qualitative information on project management and the effects of aid on innovation, partnerships, collective learning, and commercial and technological opportunities.
France	Voluntary agreement for freight companies	Reductions in CO_2 emissions Reductions in emissions of other air pollutants.	Total efficiency: total expenses (in euros) divided by the CO_2 emissions avoided (in t CO_2) Public efficiency: total public expenses (in euros) divided by the CO_2 emissions avoided (in tCO_2)	Process evaluation including an online survey of participants, interviews with various stakeholders and case studies in a sample of regions. This enabled to assess the coherence and relevance of the scheme, and to make suggestions to improve it.
Germany	Energy Efficiency Fund (data for the sub-measure "support for highly efficient cross-cutting technologies in SMEs")	Reductions in GHG emissions; Leverage effect	Funding efficiency (public funding per energy saved or emissions avoided); Administrative cost per action and per energy saved; Leverage effect (total investment per public funding)	Contribution to an energy efficient economy (qualitative, about the overall appropriateness of the measure); Satisfaction (surveys of participants and administrators); Sectoral analysis;
Germany	Energy Efficiency Networks Initiative	Reductions in GHG emissions		Satisfaction of scheme administrators and participants Statistics about the networks (type, location, runtime, etc.)

Country	Name of the measure	Other impacts monitored and/or evaluated	Indicators about (cost-)effectiveness or efficiency of the policy measure	Other indicators/aspects monitored and/or evaluated
Ireland	Better Energy Homes	Reductions in CO_2 emissions; Reductions in emissions or local air pollutants (NOx, SOx, VOCs and particulate matters); Employment effect	Net Present Value (NPV) (taking into account 3 different views: government, participants and the whole society); Cost-effectiveness indicators (in euros/kWh or tCO_2 saved) based on a Cost-Benefit Analysis	Satisfaction with the scheme and the contractors (based on a survey of participants)
Italy	White Certificates Scheme		Public expense for an additional kWh saved (until 2011); Impact of the annual cost of the scheme for a typical family (until 2011); Analysis of the ratio between the value of the certificates linked to a given project to its capital cost (mostly for large projects)	Analysis of where the certificates are obtained (per sector and/or action type); Analysis of market behaviours (e.g., capability of prioritising the most cost effective solutions)
Lithuania	Renovation programme with EU funding		average renovation cost/m^2	Technical indicators were monitored on pilot projects (indoor air temperature, relative moisture, CO_2 concentrations, heat transfer coefficient of building components) Satisfaction of flat owners and residents (survey) National Audit Report
Netherlands	Multi-year agreements in the industry		implementation costs for government and industry in relation to the benefits	Energy efficiency indicators and share of renewable energy use. Trends in the costs of the scheme. How the scheme can be improved.
Netherlands	Subsidy scheme for housing corporations in Amsterdam	Reductions in CO_2 emissions Comfort improvements Affordability of energy	cost-effectiveness in terms of public expenses per tCO_2 saved	Comfort improvements (assessed through survey of occupants); Affordability of energy (ratio of energy expenses on total income + survey of occupants)
Netherlands	Purchase tax on passenger cars	Reductions in CO_2 emissions Impact on public budget		Leakage effects of the Dutch tax measure (possible impacts on cars sold in other countries; “waterbed effect”)

Country	Name of the measure	Other impacts monitored and/or evaluated	Indicators about (cost-)effectiveness or efficiency of the policy measure	Other indicators/aspects monitored and/or evaluated
Nordic Countries	Nordsyn (market surveillance cooperation)		cost-benefit of market surveillance calculated as return on investments (costs of market surveillance activities vs. savings on electricity bills by removing non-compliant products from the market)	The evaluation made suggestions about the strategy for appliance tests.
UK	Supplier Obligations	Reductions in CO_2 emissions; Value of health benefits, air quality benefits (ex-ante);	Net Present Value (ex-ante); Value to society of lower energy bills in low income households (ex-ante)	Value of comfort taking (ex-ante); Process evaluation (suppliers' strategies, targeting of vulnerable customers, impact on the energy efficiency industry, role of public authorities)
UK	Warm Front	Reductions in CO_2 emissions; Health impacts	Value for money in comparison with the general market	Targeting of the scheme Satisfaction of the participants and comfort taking Performance of the scheme was monitored and assessed through various indicators and approaches Benefits or negative impacts on the supply chain (for ex., on turnover, jobs, etc.)
US	Auctions for capacity markets	capacity cleared and verified capacity performance		For utility EE portfolios that bid into the ISO-NE forward capacity market, they are typically also subject to other types of evaluation than the peak savings verification. While these evaluations often do not affect the ex post peak demand reduction verification, they are valuable for understanding how the design and implementation of EE schemes can be improved and forecasting market trends and system operational needs.
US	Weatherization Assistance Program	Health related impacts Quantity of avoided emissions and value of avoided emissions (state-level) macro-economic employment impacts	Savings to Investment Ratio (SIR): lifetime energy bill savings divided by the costs, both for actions done in a given year	Participant home (i.e., occupant, recipient) survey about budget issues, energy conservation behaviors, use of programmable thermostats and health issues Case studies at state's level to analyse of how WAP services are delivered + how WAP agents and clients perceive service delivery Special Technical Studies to investigate technical issues such as air sealing, duct sealing, furnace efficiency, and refrigerators Weatherization Deferral Study