Impacts and cost-effectiveness of major energy efficiency policies for existing buildings: what do we exactly know and what can we learn?

Jean-Sébastien Broc
Broc Research & Consulting
34 quai de Versailles
FR-44000 Nantes
France
jsb@brnc.fr

Catherine Guermont
ADEME, Service Climat
Angers
France
catherine.guermont-bernardi@ademe.fr

Christian Deconninck
ATEE
France
c.deconninck@atee.fr

Marie-Laure Nauleau
ADEME, Service Economie et Prospective
Paris
France
marie-laure.nauleau@ademe.fr

Keywords

cost effectiveness, existing buildings, national energy efficiency action plans (NEEAPs), data, evaluation, energy efficiency policy, comparisons

Abstract

There have been many reviews about how energy efficiency policies for existing buildings work, but more rarely about the details of their effective impacts and costs, mainly due to difficulties in accessing data. The National Energy Efficiency Action Plans enable to know what policies are implemented and how. But details about their impacts and costs can often be found in national language only, and rarely in a single report (if available at all).

This paper presents an analysis of ex-post evaluations and monitoring reports available for 13 major energy efficiency schemes for retrofitting buildings. The first part reviews the data about evaluation methods, energy savings, and costs. It describes the main data found and discusses to what extent these data can be compared. This stresses the need for in-depth study to correctly analyse the data.

The second part proposes a methodology to compare results and costs of different policies, illustrated by one pilot comparison between the Italian and the French tax credit schemes. This example shows the importance to take into account the differences in the scope and methods used to estimate the impacts and costs when comparing indicators such as cost-effectiveness ratios.

The paper concludes with a discussion about the need to promote information sharing between the Member States and evidence-based approaches for the design and management of energy efficiency policies.

Introduction

Wade and Eyre (2015) systematically reviewed the literature about evaluation of household energy efficiency programme. They analysed the main types of evaluation methods in use and the circumstances in which they are most appropriate. They also identified gaps, i.e. issues not covered by the methods available, or not addressed in the evaluations reviewed. Their study provides an overview of the state of knowledge regarding these issues. Their evidence base was restricted to peer-reviewed papers due to the scope of the study.

For this paper, we used a complementary approach, by looking at the evaluation practices used for some of the major measures (policies or programmes)[1] aiming at improving the energy efficiency of existing buildings. The scope of the study is focused on official data and evaluations, i.e. data published by the organisations in charge of the measures, by public institutions and/or in studies commissioned by public institutions or agencies.

This study was done within the French initiative EEPPEE (Evaluation of Efficiency of Public Policies on Energy Efficiency) launched by ATEE[2] in 2015 with the support of ADEME[3]. EEPPEE aims at promoting practices and uses of evaluation in order to favour evidence-based approaches in the design and management of energy efficiency policies. The starting point of this study was that very few sources gathered quantitative and detailed data about results of these policies, in particular efficiency or cost-effectiveness data (e.g., in terms of euros spent

---

[1] In this paper we use "measures" as a term encompassing both, policies and programmes.

[2] ATEE: Technical Association for Energy and Environment, gathering French energy efficiency stakeholders.

[3] ADEME: French Environment and Energy Management Agency.

Table 1. Overview of the template used for the case studies.

Analysis of the measure and its background	Details about the energy efficiency actions promoted by the measure	Analysis of the evaluation practices
> **Policy theory** (main features, how it worked) > **Background specific to the measure** (barriers to overcome, etc.) > **General background for the measure** (building stock, interactions with other measures, etc.)	> **Eligible criteria or targeted actions** (energy performance criteria, etc.) > **Eligible costs and incentive rates** > **Details about the action implementation** (requirements on installers, quality assurance processes, etc.)	> **Evaluation approach** (monitoring the actions; ex-ante/ex-post evaluation; types of evaluation methods) > **Main data collected** > **Key methodological choices** (types of baseline; adjustments applied/taken into account; causality assessment)
Results of the measure	**Direct impacts of the measure**	**Data about the costs/investments**
> **Overall participation results** (number of participants; number of actions installed; etc.) > **Results per type of action**	> **Energy savings** > **CO_2 emissions avoided** > **Other direct results evaluated** > **Information about the uncertainties**	> **Public costs/investments** > **Costs/investments by the participants** > **Costs/investments by other stakeholders** > **Costs of the actions**
Other information on results/impacts of the measure	**Analysis of the cost-effectiveness and efficiency of the measure**	**Discussions in view of a comparison**
> **Market transformation effects** > **Socio-economic impacts** > **Other impacts evaluated** (when relevant)	> **Indicators used in the communication about the measure** > **Estimation of a "€/kWh saved" ratio** (when possible/relevant) > **Other indicators available**	> **Methodological discussions** > **Aspects to be taken into account when comparing these results with other measures**

per kWh saved); and that available data were very difficult to compare from one scheme to another.

The objectives of this study were thus to review available ex-post evaluations to see what data could be gathered, how these data were obtained, if they could be compared in a consistent manner (and if yes, how), and to analyse what lessons can be learnt from this type of review and comparative approach.

After explaining the methodology, the paper provides an overview of the selected case studies, and analyses both evaluation practices and robustness of the data found. Then, the approach to compare case studies is illustrated through the comparison between the Italian and French tax credit schemes for the energy efficiency of dwellings.

Methodology used for the study

Selection of the case studies

The study was focused on policies including financial incentives or obligation schemes to improve the energy efficiency of buildings[4], assuming that these measures represent major investments from the public budget and the obligated parties. The starting point was the list of "successful policies" from the MURE database[5].

After taking into account criteria regarding the data availability and experience feedbacks (measures old enough for results to be observable), 59 cases were discussed with the members of the steering committee, and 15 cases were finally selected according to their interest: cases that can provide useful insights to the French context, having significant impacts in terms of energy savings, and forming a mix to cover both, housing and services. A first research to verify data availability led to substitute or withdraw a few cases when not enough data were found. Difficulties were encountered in particular with measures for commercial buildings.

Data sources used

The main data sources used for each case study were the MURE database, the NEEAPs (National Energy Efficiency Action Plans) and annual reports of the Member States for the EED (Energy Efficiency Directive, 2012/27/EU), websites of the organisations in charge of the measures and/or of their reporting and/or evaluation, scientific papers, and contacts with national experts. We gave priority to official data sources to analyse current evaluation practices of the public authorities, and because they are the data used for reporting at the EU level.

Scope and general framework for the analysis

The study looked at methods and data from the monitoring and ex-post evaluation of the measures. In this paper, we use "ex-post evaluation" for any study done to assess the results of a measure after its implementation (mostly impact evaluations or reviews of monitoring data) and/or to understand how the measure has worked and could be improved (mostly process evaluations or surveys of participants and/or stakeholders).

Overview of the case studies

PRESENTATION OF THE CASE STUDIES

The 13 case studies done cannot be presented in details in this paper that is why Table 2 gives only an overview providing name, country, period and general type. It also includes the cor-

---

[4] More specifically the energy consumption for space heating and domestic hot water.

[5] http://www.measures-odyssee-mure.eu/successful-measures-energy-efficiency-policy.asp

Table 2. Overview of the measures reviewed.

MURE code (country)	Name of the measure	Period	Type of measure or incentive
HOU-BEL30 (Belgium-Wallonia)	Financial incentives for energy saving investments in buildings (*Primes Energie / Réhabilitation*)	2005–on-going	Grants
HOU-FRA7 (France)	Tax credit scheme *(Crédit d'Impôt Développement Durable)*	2005–on-going	Tax credit (on personal income tax)
HOU-FRA31 (France)	Zero-Interest Rate Loan scheme *(éco-Prêt à Taux Zéro)*	2009–on-going	Soft loans
HOU-GER33 (Germany)	KfW Energy-efficient Refurbishment (KfW *Energieeffizient Sanieren Program*)	1995–on-going	Soft loans and grants
HOU-IRL42 (Ireland)	Better Energy Homes	2011–on-going	Grants
HOU-ITA30 (Italy)	Fiscal incentives for energy savings in existing buildings (*Detrazioni fiscali per il risparmio energetico del patrimonio edilizio esistente*)	2007–on-going	Tax credit (on personal or company income tax)
HOU-NLD27 (Netherlands)	Covenant energy savings rent sector (*Convenant Energiebesparing huursector*)	2008–2020	Voluntary agreement
HOU-UK20 (United Kingdom)	CERT (Carbon Emission Reduction Target)	2008–2012	Energy efficiency obligation scheme
HOU-UK5 (England)	Warm Front scheme	2000–2012	Grants
TER-AU12 (Austria)	Federal Property Contracting Programme (*Bundesimmobiliencontracting*)	2001–on-going	Energy performance contracting
TER-UK12 (United Kingdom)	CRC (Carbon Reduction Commitment) Energy Efficiency Scheme	2010–on-going	Carbon allowances
GEN-DK6 (Denmark)	Energy Saving Obligations for Energy Distributors (*Energiselskabers Energispareindsats*)	2006–on-going	Energy efficiency obligation scheme
GEN-FRA1 (France)	White certificates scheme (*Certificats d'Economies d'Energie*)	2006–on-going	Energy efficiency obligation scheme

HOU: measures for the residential sector; TER: measures for the service sector; GEN: cross-sectoral measures.

responding code used in the MURE database[6] for each measure, so that the reader can easily find more detailed descriptions.

ANALYSIS OF THE EVALUATION PRACTICES

We first reviewed what type of data sources were available about the results of the measures. We distinguished official data collected along the implementation of the measures about their direct outputs (*monitoring data*) and/or about the impacts (energy and/or CO_2 savings) but not connected with an ex-post evaluation (*estimated results*); data from official ex-post impact evaluation focused on energy and/or CO_2 savings (*impact evaluation*); data from official process evaluation (*process evaluation*); data from other evaluations, like evaluation on health impacts, job impacts, etc. (*other evaluations*); and data from scientific papers (*scientific papers*).

We then reviewed the main characteristics of the methods used for the evaluation of the energy savings. The typology used here is mainly based on the one defined in the EMEEES project (see Thomas et al., 2009). The baseline is here the situation or data used to define the baseline energy consumption when calculating the energy savings. The adjustments considered in the table below are adjustments to take into account differences between the theoretical energy consumption estimated with engineering calculations (hereafter named "conventional energy consumption") and the observed energy consumption based on metered data. In addition, we looked at whether and how causality was assessed or taken into account. In this paper, "causality" means whether the measure had a triggering or upgrading effect on the actions installed. A triggering effect means that the action would not have been installed in the absence of the measure. An upgrading effect means that the action may have been installed in the absence of the measure, but with a lower energy performance level.

Note: for the energy efficiency obligation schemes (HOU-UK20, GEN-DK6 and GEN-FRA1), only the actions on heating systems and building envelope are taken into account in the analysis of the calculation methods.

Only the measures targeting the service sector (TER-AU12 and TER-UK12) and energy efficiency obligation schemes for

---

[6] http://www.measures-odyssee-mure.eu/

Table 3. Overview of the types of sources available.

	HOU-BEL30	HOU-FRA7	HOU-FRA31	HOU-GER33	HOU-IRL42	HOU-ITA30	HOU-NLD27	HOU-UK20	HOU-UK5	TER-AU12	TER-UK12	GEN-DK6	GEN-FRA1
Monitoring data	⬜	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟩	🟩	🟩
Estimated results	🟦	🟩	⬜	🟩	🟩	🟩	🟩	🟩	🟩	🟨	🟩	🟩	🟩
Impact evaluation	⬜	🟩	⬜	🟩	🟩	🟩	⬜	🟩	🟩	⬜	🟩	🟩	⬜
Process evaluation	🟩	🟩	🟦	⬜	🟩	⬜	⬜	🟩	🟩	⬜	🟩	🟩	🟩
Other evaluations	⬜	🟩	⬜	🟩	⬜	🟩	🟦	⬜	🟩	⬜	⬜	⬜	⬜
Scientific papers	⬜	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩

⬜ no data/report found 🟩 data found 🟨 partial data found 🟦 data for a package of measures

Table 4. Main characteristics about the official evaluations of the energy savings.

MURE code (country)	**Evaluation method (energy savings)**	**Type of baseline**	**Data about energy consumption**	**Adjustments / causality**
HOU-BEL30	simplified engineering calculations (standardised values)	baseline based on national statistics and updated on a regular basis	conventional energy consumption or national statistics	no adjustment; causality partly taken into account through the definition of the baseline
HOU-FRA7	building stock modelling	scenario without the measure	conventional energy consumption	no adjustment; causality taken into account through the assumptions used in the baseline scenario
HOU-FRA31				
HOU-GER33	simplified engineering calculations on a sample of participating dwellings	energy consumption before the actions	conventional energy consumption	adjustment factor for the consumption before actions; no causality assessment
HOU-IRL42	simplified engineering calculations (+ econometric billing analysis)	energy consumption before the actions (+ ”differences in differences” method for the billing analysis)	conventional and metered energy consumption	default values for rebound and free-rider effects
HOU-ITA30	simplified engineering calculations	energy consumption before the actions	conventional energy consumption	no adjustment; no causality assessment
HOU-NLD27	monitoring of the energy performance certificates	energy consumption before the actions	conventional energy consumption	no adjustment; no causality assessment
HOU-UK20	simplified engineering calculations (standardised values)	energy consumption before the actions, or regulations according to the action type	national statistics	Default values for comfort taking; additionality requirements
HOU-UK5	simplified engineering calculations (+ billing analysis)	energy consumption before the actions (+ control group for the billing analysis)	conventional and metered energy consumption	assessment of comfort improvement; no causality assessment
TER-AU12	energy audits + M&V	energy consumption before the actions	metered energy consumption	adjustments included in energy performance contracting
TER-UK12	econometric billing analysis	“differences in differences” method	metered energy consumption	causality through the “differences in differences” method
GEN-DK6	simplified engineering calculations (standardised values or specific calculations similar to energy audits)	energy consumption before the actions or average energy consumption of equipment sold on the market	national statistics (for standardised actions) or metered energy consumption (for specific actions)	evaluation of the free-rider effects
GEN-FRA1				no adjustment; additionality requirements

the cases of large buildings (GEN-DK6 and GEN-FRA1) use evaluation methods based on metered energy consumption. The methods for the measures targeting the residential sector are either using conventional energy consumption or average values per building types from national statistics.

The methods reviewed for the residential measures are indeed all using simplified engineering calculations (directly, or within a building stock modelling for HOU-FRA7 and HOU-FRA31). This makes easier and less expensive the monitoring of the energy savings from a large number of actions. This approach also includes sometimes the implicit assumption that when the number of actions is large enough, the average result per action (for a given action type) is close to the result that would be calculated as an average of the dwelling stock (due to the Law of Large Numbers). In a few cases, this monitoring was complemented by a billing analysis on samples comparing a "participants" and a "control" group, using the statistical method of differences in differences (see Hong et al. 2006 for HOU-UK5; Scheer et al., 2013 for HOU-IRL42) or national databases about metered energy consumption (see DECC, 2014 and Adan and Fuerst, 2016 for HOU-UK20).

The preference given to simplified engineering calculations may be due to several practical reasons making the use of statistical methods difficult (Wade and Eyre, 2015): access to metered energy consumption data (e.g., energy bills) when large samples and time series are needed; forming a control group meeting the statistical requirements; collecting data about the explanatory variables simultaneously with the energy consumption data.

Even when using the same type of calculation method, the evaluations reviewed are very diverse in their methodological choices, as the definition of the baseline and the choice of using (or not) adjustments. This may be due to pragmatic choices to use the assumed best data available (e.g. if previous studies proved significant comfort taking), and/or for consistency with the policy objectives (e.g. when using current regulations as benchmark for the baseline). The causality issues are analysed further in the discussions below.

MAIN DATA FOUND

Table 5 presents a selection of the main data found related to costs (amounts of financial incentives and amounts of investments) and reported final energy savings. Investments data are mostly full costs of the actions (costs of equipment/materials, and labour costs), except for the two tax credit schemes (HOU-FRA7 and HOU-ITA30) where the data available are the eligible costs (i.e. the costs taken into account to calculate the tax credit, see details in the comparison of both schemes further on).

Data about costs have been averaged in annual terms to give an idea about the financing flows. However, there may be important variations from year to year, according to the measure. Data about energy savings are the most recent official data found to give an idea of the achievements of the measures, and also to show that these data are available for very different periods from one measure to the other. This is not only because some measures are older than others, but also because energy savings data are not always available on a systematic annual basis.

DISCUSSIONS ABOUT THE DATA FOUND

Overall results that look big, but to be considered carefully

8 of the 13 measures included public incentives. The sum of these amounts represents an overall annual average close to €5 billion/year (average value to be taken with caution due to strong variations between years).

Estimates for investments could be found for 7 measures, and amount to more than €18.5 billion/year. Again the averages calculated hide strong annual variations, and the scope of costs taken into account is different with 2 measures dealing with eligible costs and 5 measures dealing with full costs.

The achieved energy savings reported for the 13 measures would amount to about 94 TWh/year of final energy, i.e. about 1 % of the total final energy consumption of the 9 countries of these measures (8,446 TWh in 2014, Eurostat data). However, summing up these energy savings raises many questions due to the differences in the calculation methods, assumptions done, periods taken into account, etc. This is an order of magnitude that should be read with a lot of caution. Moreover, most of these values include a high level of uncertainty.

Comparisons between measures, when possible, require an extreme caution

The data in Table 5 cannot be compared directly, because:

1. The magnitude of the measures is dependant of the context (e.g., size of the building stock). This could be taken into account by calculating indicators in relative terms (e.g., € of financial incentives per € of GDP).
2. The scope of the measure may be very different (for example, HOU-UK5 targeted fuel poor, not all households). This could be taken into account by using ratios relating the results of the measure to the number of dwellings in the scope of the measure. However, this scope is not always clearly defined.
3. The periods are different. This can be addressed by calculated annual averages. But these annual averages have to be taken with caution, due to the sometimes strong variations between years.
4. The scope of the results may be different (e.g., for the data about investments, see comparison below).
5. The methods used to estimate the results differ. Even if the energy savings can easily be expressed in the same energy unit (e.g. TWh of final energy), they do not represent comparable quantities, due to differences in the methodological choices (see above, and discussions about "net" and "gross" results below). The data found were not detailed enough to allow to re-compute the energy savings in a harmonised way.

Possible to assess the effectiveness of the measures, but with an unknown uncertainty

We distinguish in this paper "gross" and "net" results, as well as three evaluation indicators (effectiveness and efficiency of the measures, and cost-effectiveness of the energy efficiency actions), as defined in the table below.

The data found about achieved/reported energy savings and outputs enable the assessment of the effectiveness of the meas-

Table 5. Main data found.

MURE code (country)	Amounts of financial incentives	Amounts of investments	Reported final energy savings
HOU-BEL30	M€44/y (average over 2010–2012)	No official data found	926 GWh/y from actions over 2009 2012 (SPW, 2014)
HOU-FRA7	M€1,600/y (average 2006–2014)	M€6,300/y (average over 2006–2014) (eligible costs, see details in the comparison below)	14.9 TWh/y from actions over 2005–2012 (NEEAP2011)
HOU-FRA31	M€117/y (average over 2009–2014)	M€854/y (average over 2009–2015) (full costs of actions)	2.1 TWh/y from actions over 2009–2012 (NEEAP2014)
HOU-GER33	M€800/y (average over 2009–2015)	M€6,400/y (in 2015) (full costs of actions)	15.9 TWh/y from actions over 2005–2014 (IWU and IFAM, 2015)
HOU-IRL42	M€26/y (average over 2009–2015)	M€87/y (average over 2009–2016) (full costs of actions)	312 GWh/y for actions over 2008–2010 (Scheer and Motherway, 2011)
HOU-ITA30	M€1,900/y (average over 2008–2014)	M€3,400/y (average over 2008–2014) (eligible costs, see details in the comparison below)	11.2 TWh/y for actions over 2007–2014 (2016 annual report for the EED)
HOU-NLD27	M€100/y (average over 2014–2017)	No official data found	2.3 TWh/y for actions over 2011–2014 (RVO, 2015)
HOU-UK20	(1)	M€1,250/y (average over 2008–2012) (full costs of actions)	9 TWh/y for actions over 2010–2012 (NEEAP2014)
HOU-UK5	M€230/y (average over 2000–2012)	More than M€230/y (average over 2000–2012) (3)	(2)
TER-AU12	*not applicable*	No official data found	73 GWh/year for actions over 2001–2013 (NEEAP, 2014)
TER-UK12	(5)	No official data found	(4)
GEN-DK6	No official data found	No official data found	8 TWh/y for actions over 2008–2012 (NEEAP2014) (6)
GEN-FRA1	No official data found	No official data found	29 TWh/y for actions over 2006–2013 (NEEAP2014) (6)

(1) About M€1,000/y (average over 2008–2012) of costs reported by the obligated parties.

(2) No comprehensive data found about the part of investments paid by owners or third party.

(3) While the NEEAP2011 reported 8.0 TWh/year (from actions over 2000–2010), the NEEAP2014 (p. 134) stated that "the latest estimates of energy savings from this policy show negligible changes in energy consumption due to high comfort taking". This may be due to the difference between the engineering calculations (2011) and the billing analysis (2014), particularly important for a measure focused on fuel poverty.

(4) About M€700/y of public revenues from the carbon allowances (vs. less than M€40 administration costs).

(5) Annual average reduction of more than 10 % in gas consumption and 3–5 % in electricity consumption between 2010 and 2012 for the participants compared to the control group (CAG Consultant et al., 2015).

(6) For actions from all sectors (not only for buildings).

Table 6. Main evaluation indicators considered in this paper.

Indicator	Definition used in this paper
"gross" results	all results monitored for the measure (e.g., all actions receiving a grant from the measure)
"net" results	results that can be attributed to the measure after a causality assessment and/or by taking into account additionality criteria in the baseline (e.g. energy savings taking into account free-rider effects)
Effectiveness of the measure	**"gross" achievements** of the measure related to its **objectives** (e.g., % of achievement of the target number of dwellings renovated)
Efficiency of the policy measure	**costs of the measure** (i.e. public expenditures and/or costs for the obligated parties, including administration, marketing, incentives and other costs) related to its **"net" results** (e.g. public cost of "net" avoided CO_2 emissions in €/tCO_2 avoided)
Cost-effectiveness of the action	**costs of the actions** related to their **"gross" results** (energy or CO_2 savings) (usually in €/kWh saved or in €/tCO_2 avoided)

ures (when quantitative targets were set). While the outputs are usually monitored with a good level of accuracy, the uncertainty level of the reported energy savings is more rarely discussed. The studies that investigated further this issue found important sources of uncertainty (Adan and Fuerst, 2016; DGFIP, 2011; Hong et al., 2006; Scheer et al., 2013). In most cases using simplified engineering calculations, the reported "gross" energy savings are very likely overestimated, in particular due to overestimation of energy consumption before actions (prebound effect) and underestimation of energy consumption after actions (due to rebound effect and performance gaps) (see for example, Galvin, 2014). But it is currently difficult to know to what extent.

Some of the measures using simplified engineering calculations have then used correction factors to account for these issues (see Table 4), using default values defined from available studies.

The evaluation based on billing analysis (CAG Consultants et al., 2015; Hong et al., 2006; Scheer et al., 2013) includes confidence intervals estimated with statistical methods. However these confidence intervals do not encompass all the sources of uncertainty, for example due to bias in sampling or matching samples.

Causality, additionality, "gross" and "net" results

An official study looking at the causality of the impacts could be found for very few measures. This was done either by simulating "without" and "with measure" scenarios and comparing different modelling approaches (DGFIP, 2011 for HOU-FRA7) or by using the statistical differences-in-differences method (CAG Consultants et al., 2015 for TER-UK12; Hong et al., 2006 for HOU-UK5; Scheer et al., 2013 for HOU-IRL42). However, except for TER-UK12, the results from these studies have not been used in the official communication about these measures (for example for the results reported in the NEEAPs).

For some measures (GEN-DK6, HOU-FRA7, HOU-FRA31 and HOU-IRL42), the official calculation methods include default values or assumptions to take into account a share of free-riders (participants who would have done the actions anyway). Another approach used to tackle the causality issue is to use additionality criteria in the definition of the baseline (HOU-BEL30 and HOU-UK20). For example, by defining the baseline as the average energy efficiency of the equipment sold (or as the minimum energy efficiency standards), and not as the average energy efficiency of the equipment replaced. This means assuming that the equipment would have been replaced anyway, and implicitly that the measure has an upgrading effect but no significant triggering effect.

The results of energy savings for these measures (GEN-DK6, HOU-BEL30, HOU-FRA7, HOU-FRA31, HOU-IRL42, HOU-UK20 and TER-UK12) can thus be read as "net" results. However, the extent to which the causality has been taken into account varies strongly from one measure to the other. And none but one (TER-UK12) of the official results reported for these measures includes an ex-post assessment of the triggering and/or upgrading effects. The results for the other measures should be read as "gross" results, i.e. calculated with the implicit assumption that all actions counted for the measures would not have been implemented in their absence.

Studies done independently by researchers investigated the issue of free-riders (Alberini et al., 2014 about HOU-ITA30; Grösche et al., 2013 about HOU-GER33; Nauleau, 2014 about HOU-FRA7). They showed sources of significant free-rider effects, but their results could not be extrapolated due to data limitations.

In parallel, as pointed out by Wade and Eyre (2015), we did not find any official quantitative assessment about possible non-participant spill-over effects. Several evaluation reports highlighted that the requirements for the financial incentives may have driven market transformations towards higher market shares for the most energy efficient products or solutions, and/or towards better quality. But the data collected did not make possible to go beyond qualitative appraisals. This is an important point when considering the "net" results of a measure: while the free-rider effects may reduce its impacts on short term, the spill-over effects may increase them on a longer term (as mentioned by IWU and IFAM (2015) for the case of HOU-GER33).

Very few official indicators about the efficiency of the measures

Official efficiency ratios were found only in (DGFIP, 2011) for HOU-FRA7 (see comparison below) and in (Scheer and Motherway, 2011) for HOU-IRL42 (see below). For all the reasons mentioned above about the risks of direct comparisons, we chose not to calculate directly efficiency ratios from the data found. Their interpretation in a comparative purpose does require a detailed analysis (see next part).

Scheer and Motherway (2011) made an evaluation of the efficiency of HOU-IRL42 by estimating the Net Present Value (NPV)[7] of the actions receiving a grant, taking into account free-rider effects through a default value. They computed NPV ratios according to different viewpoints (public budget, society) and scenarios (for energy prices and types of externalities taken into account). Scheer and Motherway pointed that, despite all their efforts 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 (factor 4 on the Net Present Values between the "low energy prices" scenario and the "high energy prices" scenario when taking into account the energy savings only). All the scenarios computed for the "society" NPV showed a positive result (i.e. a benefit to the economy) (0.024 €/kWh for the "medium" scenario of energy prices when not taking into account positive externalities).

Scheer and Motherway (2011), as well as DGFIP (2011), also warned the readers that classical efficiency ratios (e.g., € invested/kWh saved) only provide part of the information about the impacts and efficiency of the measures. Their long term effects and non-energy benefits (for example, on health or employment) are often difficult to assess, and even more to monetize. But they should not be disregarded. First, because energy effi-

---

[7] Defined in their study as the present value of an investment's future net cash-flows minus the initial investment, and taking into account the time value of money by applying a discount rate to future cash-flows.

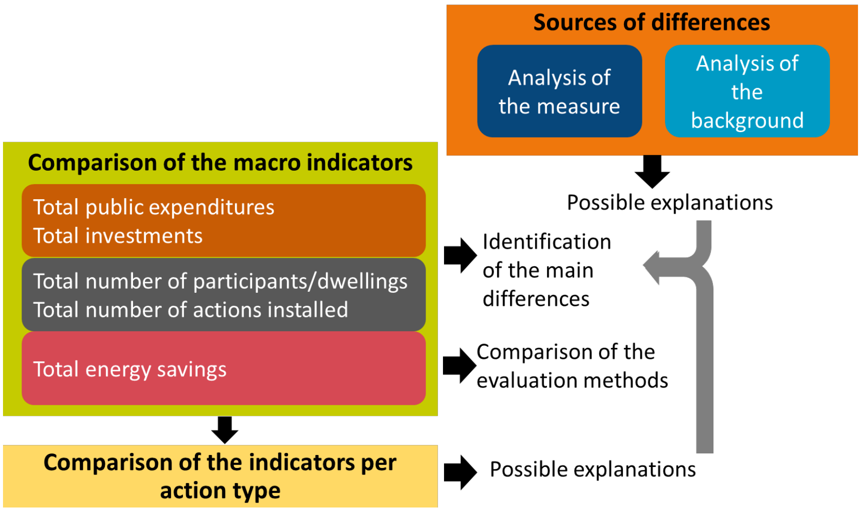

Figure 1. Overview of the comparative approach.

ciency policies have often multiple objectives and should therefore be evaluated through multi-criteria approaches. Second, because non-energy benefits may be of upmost importance in the decision making of participants (for example, comfort improvement) and of policy makers (for example, effects on public budget and employment, as analysed for HOU-GER33 by Rosenow (2013)).

Comparative approach

METHODOLOGY USED

Our comparative approach (summarised in Figure 1) starts by comparing the macro indicators (i.e. for the whole measure) to identify the main differences and explain them by further analysing the same indicators per action type, the main characteristics of the measures (eligibility criteria, incentive rates, etc.) and the background (characteristics of the dwelling stock, interactions with other measures, etc.).

The approach was tested on measures implemented for 10 or more years taking into account changes made to the measures. It would have been interesting to also analyse market trends. But only limited data could be found.

We illustrate below the approach through a pilot case, focusing on main results and methodological discussions.

PILOT CASE: COMPARISON OF THE ITALIAN AND FRENCH TAX CREDIT SCHEMES

Both schemes provide tax credits on the income tax for households[8] who invested in energy efficiency actions meeting the required eligibility criteria. Both schemes were created in continuation of tax credit schemes for renovation actions that did not include any energy efficiency criteria. In France the new scheme (with energy efficiency criteria) replaced the old one. In Italy both schemes (without and with energy efficiency criteria) co-exist, but the "energy efficiency" scheme has a higher incentive rate.

Comparison of the macro-indicators

Both schemes have similar magnitude of public expenditures, however they represented a higher % of GDP in Italy (lower GDP in Italy, and more impacted by 2008 crisis). The French scheme delivered a much higher renovation rate for similar budgets, because of a lower incentive rate (on average: 28 % vs. 57 % in Italy) and a smaller scope of costs eligible (equipment/materials costs only in France[9] vs. full action costs in Italy, including costs of the actions directly induced by the action). Figures 2[10] and 3[11], [12] show the evolutions over time.

The Italian incentive rate was simpler (one single rate) and remained stable (only one change in June 2013, from 55 to 65 %), whereas the French rates were more complex (several rates according to the action types). However, the French modalities didn't create a barrier up to 2011, as the participation was high (see figures below). The drops in participation in 2011 and 2012 were mainly due to a decrease in the incentive rates and to the introduction of the concept of "bunch of works" that aimed at stimulating projects with higher ambitions. The decrease in the incentive rates in France was decided due to the context of public debt crisis. At the opposite, the incentive rate was maintained in Italy, and even increased in 2013, as part of the stimulus plan.

The qualitative cross-analysis over time of the public expenditures and outputs with background factors (e.g., trends in GDP and energy prices) did not make possible to draw any conclusion at the level of the macro indicators. The changes in the measures' characteristics (in particular the incentive rates) seemed to have a higher influence on the participation than the background factors.

---

[8] Also for legal entities in Italy, but 95 % of the applications have come from households.

[9] Except for insulation of walls and roofs (with installation costs eligible).

[10] Tax credit claimed: In practice, the public expenditures occur the year after (for France) or over several years (for Italy). Therefore the amounts of tax credit claimed each year to enable a direct comparison.

[11] The action types eligible for the French scheme also include actions that are not energy efficiency actions nor equipment for renewable heat. These include for example PV panels up to 2013, or voluntary Energy Performance Certificates from 2009. These actions are not included here for the comparison between the Italian and French schemes to be consistent (same scope of action types eligible).

[12] DGFIP annual data: http://www2.impots.gouv.fr/documentation/statistiques/2042_nat/Impot_sur_le_revenu.htm.

Table 7. Macro-indicators for the Italian and French tax credit scheme.

Country	Public expenditures	Outputs	Rate of final energy savings*
Italy	About €1.9 billion/year (about 0.12 % of GDP) (average over 2008–2014)	About 300,000 dwellings/year (renovation rate: 1.3 %/year) (average over 2008–2014)	0.13 Mtep/year ("gross" result) (0.6 % of heating consumption) (average over 2008–2013)
France	About €1.6 billion/year (about 0.09 % of GDP) (average over 2006–2014)	About 1.1 million dwellings/year (renovation rate: 4.8 %/year) (average over 2006–2014)	0.16 Mtep/year ("net" result) (0.6 % of heating consumption) (average over 2005–2012)

** New annual energy savings for the actions installed each year.*

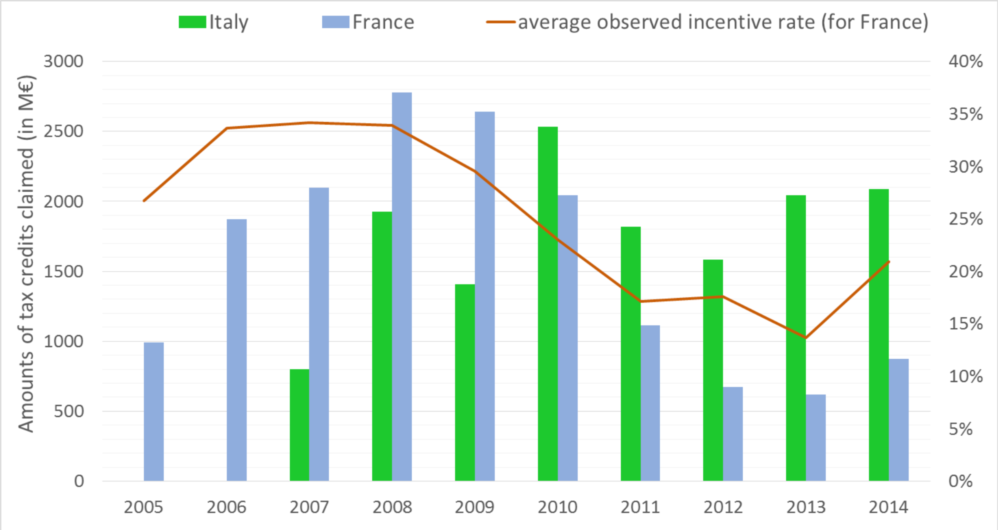

Figure 2. Annual amounts of tax credit claimed for both schemes (in million euros). Source: Annual reports by ENEA for Italy (ENEA, 2015) and annexes of the annual Budget Law for France.

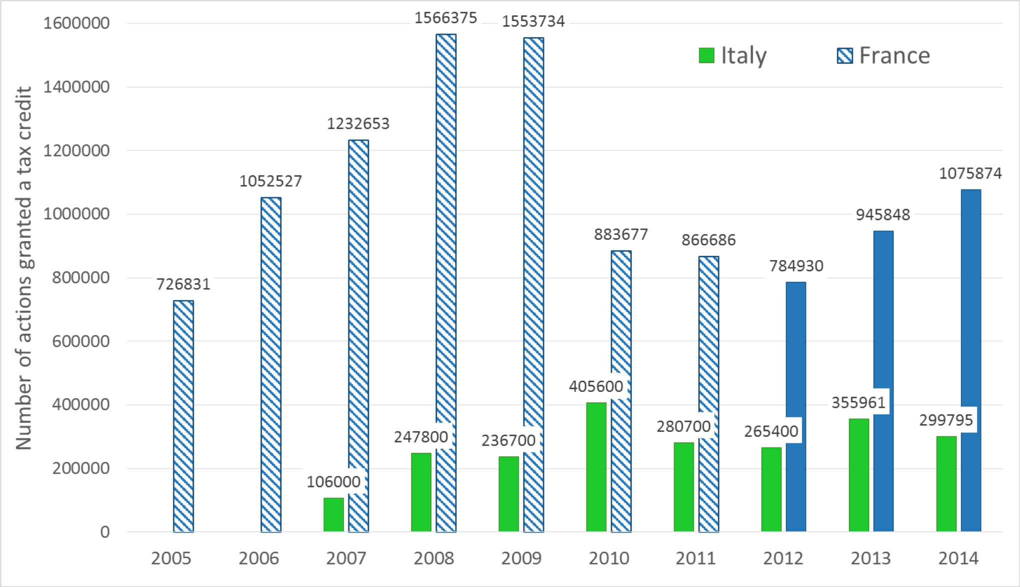

Figure 3. Number of actions granted a tax credit each year for both schemes. Note: For the French scheme, the disaggregation per action type had to be assessed in (DGFIP, 2011) for data over 2005–2011 (data in "hatched blue"). Disaggregated data were directly monitored from 2012 (data in "filled blue"). Source: Annual reports by ENEA for Italy (ENEA, 2015), (DGFIP, 2011) and annual data published by DGFIP for France.

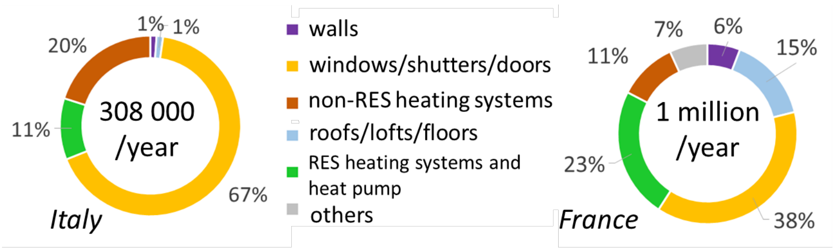

Figure 4. Distribution of the actions per type (for actions installed over 2012–2014).

Table 8. Cost-effectiveness indicator per action type for both schemes.

Scheme	Cost-effectiveness indicator / Action type	Insulation of walls	Insulation of roofs	Replacement of Windows	Solar heat systems	Condensing boilers	Heat pumps
Italy (ENEA, 2015)	full investment cost (€) per kWh saved	0.098	0.085	0.146	0.069	0.133	0.158
France (DGFIP, 2011)	marginal CO_2 abatement cost (in €/ tCO_2 avoided)	-166	-166	[-110; 387]	[740; 914]	[-238; 77]	[30; 185]

Note 1: For French results, upper bounds were calculated using full investment costs, while lower bounds used extra investment costs, i.e. when assuming that the equipment was replaced at its end of life.

Note 2: Negative values means that the cumulated and discounted energy savings overpass the investment costs.

About energy savings, the results for both schemes look similar at first sight. But they cannot be compared as such due to many differences in the respective evaluation methods. The main difference lies in the choice of the baseline: “energy consumption before actions” (Italy) vs. “scenario without the measure” (France). The French baseline scenario includes assumptions about the rate of actions that would have been installed in the absence of measure, while the Italian calculations are equivalent to assuming that all actions are attributed to the measure. The French energy savings can be considered “net” savings, and the Italian energy savings “gross” savings.

Comparison of indicators per action type

The distribution of the actions per type shows a similar hierarchy in the action types: 1) windows, 2) heating systems (RES – Renewable Energy Sources, and non-RES) and 3) insulation of opaque surfaces (walls, roofs, lofts, floors). However, the shares are different with a much higher share for windows and a much smaller share for insulation of opaque surfaces in Italy. This may be explained by a smaller share of individual houses in the Italian dwelling stock (27 % vs. 57 % in France)[13], also found in the shares of participating dwellings (40 % in Italy vs. more than 80 % in France). Italy has a higher share of condominiums (co-ownerships) where replacing windows is simpler to undertake compared to actions requiring a collective decision.

One would expect to see lower energy savings for the Italian scheme due to the lower number of actions and the higher share of windows (that deliver less energy savings than the insulation of opaque surfaces in theory). The Italian results show that the heating systems represented the highest share of the total energy savings for this scheme. This may be due to the difference in the baselines (see above): while the baseline chosen for boiler replacement in Italy is the average efficiency of the boilers in the stock, the baseline for the same action in France is equivalent to a mix of the estimated stock average (for the share of boilers replacement assumed to be anticipated due to the tax credit) and of the estimated market average (for the share of boiler replacements assumed to be done at end of life).

Evaluations of both schemes (DGFIP, 2011 for France and ENEA, 2015 for Italy) include estimates of cost-effectiveness ratios per action type (while global efficiency ratios were found only for the French case) – see Table 8[14]. However, these ratios represent distinct indicators, and the data found did not enable to compute harmonised indicators. For Italy, the indicator is a direct ratio of the investment (full cost of actions) divided by the primary energy savings over the action lifetime. For France, the indicator is the marginal CO_2 abatement cost, i.e. the marginal cost of the action divided by the CO_2 savings over the action lifetime. In (DGFIP, 2011), the marginal cost is the investment costs less the energy savings (in euros) cumulated and discounted over the action lifetime. The energy savings are calculated using the same assumptions for energy prices over time and for the discount rate (4 %/year) in the “baseline” and “measure” scenarios.

---

[13] See data in national renovation strategies: http://ec.europa.eu/energy/en/topics/energy-efficiency-directive/buildings-under-eed.

[14] Note 1: The extra investment cost is then calculated as the difference in investment costs between a “standard” equipment and a “high efficiency” equipment.

The insulation of walls and roofs are among the most cost-effective actions for both schemes, while they represent the least frequent actions (except for roof insulation in France). At the opposite, the replacement of windows (when full investment costs are taken into account) appears among the least cost-effective actions, while it represents the most frequent action. This would support the assumption that a significant share of these windows would have been replaced anyway, as suggested for example by Nauleau (2014). Alberini et al. (2014) pointed that the free-rider share would be even higher for boiler replacement, which may also explain why heating systems are the second most frequent actions despite not ranking in the most cost-effective action types (when full investment costs taken into account). Both evaluations (DGFIP, 2011 and ENEA, 2015) therefore raised the issue of the targeting of the measure. In practice, the fact that the replacement of windows or heating systems may be frequently done anyway is partly taken into account in the French scheme through the difference in scope of eligible costs between these actions (eligible costs = material/equipment costs only) and the insulation actions for opaque surfaces (eligible costs = material/equipment costs + installation/labour costs).

When comparing the cost-effectiveness ranking for both schemes, a difference stands out for solar heat systems. In addition to the difference in potential for solar energy between both countries, this may be explained by the particular French context where the assumption for the baseline takes into account that solar water heaters may replace servo electric water heaters working at off-peak hours (thereby with very low CO_2 emissions).

DISCUSSIONS

The above summary of the pilot comparison shows how risky too quick comparisons of macro results may be. A full understanding of the results from different measures requires entering into the details of their characteristics and background. Despite detailed evaluation reports (DGFIP, 2011; ENEA, 2015), the information found do not allow to re-compute data in a harmonised way to obtain comparable indicators because of data limitations: data not available (for example cost data per cost category), data not detailed in the report (for example quantitative factors transcribing the key assumptions in the energy savings calculations), or data whose scope is not explicit in the reports (for example about amounts of investments).

Conclusions and perspectives

The review of 13 major measures for improving energy efficiency in existing buildings showed that it is possible to find data sources from official monitoring and/or evaluation for each measure, but with very heterogeneous levels of details. This impeded the assessment of harmonised indicators based on the data available. Such calculations would require making default assumptions that would add uncertainty to results that already include various sources of uncertainty. The relevance of such calculations would thus be questionable.

Most of the measures reviewed are 10 or more years old. Difficulties were sometimes encountered to gather consistent data to cover the whole duration of a given measure, especially when its characteristics have changed during this period, when data are reported in different ways or sources, and that their scope is not mentioned.

One of the objectives of the study was to deeply look at the usual indicators and data available to assess the effectiveness and efficiency of the selected measures, and the cost-effectiveness of the actions. The data found made possible to assess the effectiveness of the measures, but the uncertainty about the reported energy savings raised questions that would require further investigations.

Another observation is that official efficiency indicators have been found in only 2 cases. This may be explained by the difficulty to evaluate the causality between incentives and actions, in particular when there are interactions between incentives aiming at triggering energy efficiency actions. In addition, cost-effectiveness or efficiency indicators may be seen as too restrictive for policies having multiple objectives (e.g. employment in the building industry). Still, cost-effectiveness and efficiency indicators are useful tools to improve the design, and in particular the targeting, of the measures. When cost-effectiveness indicators were found, they indeed raise questions about this (see Italy-France comparison).

Since the comparison of results between measures is very tricky, the identification of differences in the results helps targeting the analysis, when looking for their possible explanations in the characteristics and background of the measures.

The recent years have seen energy efficiency targets being increased in many countries and calls for more investments, in particular in energy renovation of buildings. A better knowledge of the actual impacts of measures and actions is therefore important to improve measures design, and explain possible gaps between expected and actual energy and CO_2 savings. Sources found during this study show that this investigation field is in development. The roll-out of smart meters may help in collecting more detailed energy data but will not solve by itself all the difficulties especially those regarding methods and procedures for data collection, treatment and analysis (for example, bias in sampling, matching of energy consumption data with data for explanatory variables, etc.). Not to mention that smart meters will not cover consumption of heating oil, biomass, etc.

Moreover, impact evaluations should not be considered separately. Impact and process evaluations should be more connected, as they inform each other: without an impact evaluation, the conclusion from a process evaluation remains qualitative; without a process evaluation, the impacts may be misinterpreted. Such combination of impact and process evaluation was rarely observed in the information found for this study.

Further improvement of evaluation practices could be the combination of long term market data and evaluation data by sector in order to take into account sectors' specificities and impacts of the different measures linked to a common market segment. Example of such approach was not found in official sources for the cases reviewed.

Last but not least, the difficulties encountered in gathering data show the need for more transparency and shared methods in the publications of results. A good practice charter could be useful to promote the publication of data in a common, robust and transparent manner, and therefore to gain confidence of stakeholders and stimulate investment by companies and consumers.

References

Adan, H., Fuerst, F., 2016. Do energy efficiency measures really reduce household energy consumption? A difference-in-difference analysis. *Energy Efficiency*, 9 (5), 1207–1219.

Alberini, A., Bigano, A. & Boeri, M., 2014. Looking for Freeriding Opportunities: Energy Efficiency Incentives and Italian Homeowners. *Energy Efficiency*, 7 (4), 571–590.

CAG Consultants, Carbon Trust, Databuild, Imperial College Business School, 2015. CRC Energy Efficiency Scheme Evaluation. Final report for the DECC, July 2015.

DECC (Department of Energy and Climate Change), 2014. Summary of analysis using the National Energy Efficiency Data-Framework (NEED). 26 June 2014.

DGFIP (Inspectorate-General for Finance), 2011. Rapport pour le comité d'évaluation des dépenses fiscales et des niches sociales: Annexe D, Synthèse de l'évaluation du Crédit d'Impôt Développement Durable. Interministerial evaluation report, April 2011.

ENEA, 2015. Le detrazioni fiscali del 55–65 % per la riqualificazione energetica del patrimonio edilizio esistente nel 2013.

Galvin, R., 2014. Why German homeowners are reluctant to retrofit. *Building Research & Information*, 42 (4), 398–408.

Grösche, P., Schmidt, C. M., & Vance, C., 2013. Identifying free-riding in home renovation programs using revealed preference data. *Jahrbücher für Nationalökonomie und Statistik*, 233 (5-6), 600–618.

IWU, IFAM, 2015. Monitoring der KfW-Programme „Energieeffizient Sanieren" und „Energieeffizient Bauen" 2014. Evaluation report by IWU and Fraunhofer IFAM for KfW. 24 November 2015.

Hong, S. H., Oreszczyn, T., Ridley, I., Warm Front Study Group, 2006. The impact of energy efficient refurbishment on the space heating fuel consumption in English dwellings. *Energy and Buildings*, 38 (10), 1171–1181.

Nauleau, M-L., 2014. Inertia And Effectiveness Of Tax Credits For Home Insulation: a French Case Study. Proceedings of IEPPEC 2014.

Rosenow, J., 2013. The politics of the German CO_2-Building Rehabilitation Programme. Energy Efficiency, 6 (2), 219–238.

RVO, 2015. Monitor Energiebesparing Gebouwde omgeving 2014. December 2015.

Scheer, J., Clancy, M., Hógáin, S. N., 2013. Quantification of energy savings from Ireland's Home Energy Saving scheme: an ex post billing analysis. *Energy Efficiency*, 6 (1), 35–48.

Scheer, J., Motherway, B., 2011. Economic Analysis of Residential and Small-Business Energy Efficiency Improvements. Report of the SEAI, September 2011.

SPW (Public Service of Wallonia), 2014. Evaluation énergétique et environnementale de la 1^e Alliance Emploi Environnement. Presentation for the Steering Committee, 2 April 2014.

Thomas, S., Boonekamp, P., Vreuls, H., Broc, J., Bosseboeuf, D., Lapillonne, B., Labanca, N., 2009. How much energy saving is 1 % per year? We still don't know but we know better how to find out. Proceedings of the eceee 2009 Summer Study, paper 3-170, 589–600.

Wade, J., Eyre, N., 2015. Evaluation good practice: is 'good enough' better than 'perfect'? Proceedings of the eceee 2015 Summer Study, paper 8-099-15, 1757–1766.

Acknowledgements

The authors wish to thank the steering committee of the study for their inputs, and Johannes Thoma, Dominique Osso and the anonymous reviewers for their useful comments.