

Theory-based policy evaluation of 20 energy efficiency instruments

Mirjam Harmelink · Lars Nilsson ·
Robert Harmsen

Received: 19 October 2007 / Accepted: 16 March 2008 / Published online: 1 July 2008
© Springer Science + Business Media B.V. 2008

Abstract Realizing a 20% energy efficiency improvement in Europe by 2020 requires the introduction of good new energy efficiency policies as well as strengthening and enforcing the existing policies. This raises the question: what characterizes good and effective energy efficiency policies and their implementation? Systematic ex post evaluation of energy efficiency policies can reveal factors determining not only what works and what does not but also explain why. Ex post evaluation of 20 energy efficiency policy instruments applied across different sectors and countries in Europe among others showed that ex post evaluation does not yet have a high priority among policy makers: Often, quantitative targets and clear timeframes are lacking, and monitoring information is not collected on a regular basis. Our analysis,

however, did reveal some general factors in the process of design and implementation of policy instruments that appear as important including (1) existence of clear goals and a mandate for the implementing organization, (2) the ability to balance and combine flexibility and continuity, (3) the involvement of stakeholders, and (4) the ability to adapt to and integrate adjacent policies or develop consistent policy packages. The analysis was performed using a uniform methodology called “theory-based policy evaluation”. The general principle behind this approach is that a likely theory is drawn up on the program’s various steps of logic of intervention to achieve its targeted impact in terms of energy efficiency improvement. The approach has several benefits over other ex post evaluation methods because (1) the whole policy implementation process is evaluated and the focus is not just on the final impacts, (2) through the development of indicators for each step in the implementation process, the “successes and failures” can be determined to the greatest extent possible, and (3) by applying this approach, we not only learn whether policies are successful or not but also why they succeeded or failed and how they can be improved.

M. Harmelink (✉) · R. Harmsen
Ecofys Netherlands,
P.O. Box 8408, 3503 RK Utrecht,
The Netherlands
e-mail: m.harmelink@ecofys.nl

R. Harmsen
e-mail: r.harmsen@ecofys.nl

L. Nilsson
Environment and Energy Systems Studies,
Lund University,
P.O. Box 118, 221 00 Lund, Sweden
e-mail: lars_j.nilsson@miljo.lth.se

Keywords Ex post evaluation · Theory-based policy evaluation · Energy efficiency · Policies

Introduction

In October 2006, the European Commission published the Action Plan on Energy Efficiency. The plan aims to realize a 20% energy efficiency improvement by 2020 (EC 2006a). This objective was reinforced in January 2007 when the European Commission presented her energy package for the twenty-first century. The package included a 20% emissions reduction target for greenhouse gases by 2020, among others, to be achieved by improving energy efficiency by 20% (EC 2007). Realizing this potential requires the introduction of good new energy efficiency policies as well as strengthening and enforcing existing policies on the European and the national level. This raises the question: what characterizes good and effective energy efficiency policies and their implementation? Systematic ex post evaluation of energy efficiency policies can reveal factors that explain success and failure related to policy design as well as the implementation process. However, while an increasing number of energy efficiency policy instruments are being introduced in Europe and elsewhere, only few of the instruments are evaluated systematically. Evaluations are often hard to compare due to the diversity in methods and indicators used. Furthermore, methods currently applied in Europe in ex post policy evaluation are mostly focused on ‘final effects,’ i.e., energy savings and cost effectiveness. Limited research is focused on bringing policy evaluation methods on an equal footing and on systematically assessing successes and failures in the implementation of energy efficiency policies. Efforts so far include:

- The SAVE project entitled “A European Ex Post Evaluation Guidebook for Demand Side Management (DSM) and Energy Efficiency (EE) Service Programmes” (SRC 2001), which provides general guidelines for ex post evaluation of DSM and EE services. These guidelines were tested for a number of DSM and EE service programs in the European Union.
- The evaluation guidebook “Evaluating Energy Efficiency Policy Measures and DSM Programmes” (IEA 2005) published by IEA DSM IA (task 1). They provide guidance for systematically evaluating the implementation process of energy efficiency policy instruments. The developed

method was applied for various types of instruments implemented in IEA countries.

Broader experience in systematic ex post evaluations that are not just focusing on “final effects” can be found in the USA. For example:

- “California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals” (Hall et al. 2006). This guidebook is to be used in evaluating California’s energy efficiency programs. The guidebook states that an evaluation plan shall at least include a brief description of the program(s) being evaluated including a high-level presentation of the program theory. The Protocols are the primary guidance tools for policy makers and evaluation contractors and has been applied various times to assess the impact, and success and failure of policies.
- Evaluations performed within the framework of the New York Energy Smart Programme (NYSERDA 2008). Examination of causality is an important component of the New York Energy Smart evaluation efforts. It is aimed at helping to validate program interventions, justify the expenditure of public benefits funds, and assist policy makers in decisions regarding future funding of public benefit programs.
- Handbook prepared for the US Department of Energy “Impact Evaluation Framework for Technology Deployment Programs” (Reed et al. 2007). The handbook provides an approach for quantifying retrospective energy savings, clean energy advances, and market effects of technology deployment programs, which includes preparing a program theory.

Within the EU-funded project “Active Implementation of the European Directive on Energy Efficiency” (AID-EE¹), we tried to contribute to the further development of a generic framework for ex post evaluations within Europe (Joosen and Harmelink 2006). The method is based on the theory-based policy evaluation as described by Blumstein et al. (2000). The method is designed to systematically

¹ Active Implementation of the Directive on Energy Efficiency (AID-EE)

assess all steps of the policy implementation process with the aim to determine success and failure factors but also final effects such as target achievement, energy savings impact and cost effectiveness, as well as success and fail factors. By developing this method, we aim to contribute to the further development and harmonization of ex post policy evaluation methods and to creating comparable evaluation outcomes. Within the project, the method was applied to evaluate 20 energy efficiency policy instruments applied across Europe, Japan, and the USA (Ecofys et al. 2007).

This paper starts with an overview of methods applied in ex post policy evaluation, followed by a description of the theory-based policy evaluation method and the practical approach developed within the AID-EE Project. Next, overall findings are presented of the 20 case studies with respect to target achievement, energy savings impact, cost effectiveness, and typical success factors per instrument. This paper does not provide the detailed results from the various case studies. These can be found on the project website <http://www.aid-ee.org> and are also presented in (Harmsen et al. 2007; Bongardt and Kebeck 2007; Khan and Nordqvist 2007; Nordqvist 2007). Finally, conclusions and a discussion are presented on the practical application of the method for policy makers.

Methods applied in ex post policy evaluation

A variety of methods are applied in ex post evaluation of policy instruments. One is the assessment of aggregate ‘top-down’ indicators on energy consumption per sector or end user. Based on statistics, a hypothetical baseline is constructed assuming that energy efficiency stays unchanged from the base year (frozen energy efficiency) or is adjusted for autonomous efficiency improvements. The actual energy use is subtracted from this amount, and the difference is defined as the amount of energy saved. This method generally does not provide much insight in the impact of individual policies due to the aggregated level of analysis.

Another approach is econometric modeling of the impact of policy instruments. In econometric modeling, a list of factors (one of which is the analyzed policy instrument) is drawn up that potentially could affect (specific) energy use of a sector. Through

statistical methods, the impact of the analyzed policy instruments can be estimated. The approach is suitable for evaluating general policy instruments, such as taxes. These methods, however, do not provide insight in ‘why’ an instrument performed or did not perform as expected and what could be done to improve it.

A third approach is detailed bottom-up policy evaluation that focuses on determining the ‘final effects’ of a single policy instrument or a package of instruments. A bottom-up calculation method means that energy savings obtained through the implementation of a specific type of energy efficiency improvement measure (e.g., a CFL) are determined in GJ or kWh that can be attributed to specific energy efficiency policy instruments. However, it must be noted that it is often difficult to determine the effect of single-policy instruments because of the fact that energy efficiency instruments often come in a package, that autonomous efficiency improvement cannot be clearly separated from policy-induced savings, and that the impact of rebound effect, free riders, etc. are insufficiently known. A combination of top-down and bottom-up evaluation methods will be the officially applied methods to evaluate the EU Energy End-use Efficiency and Energy Services Directive (ES Directive; EC 2006b).

A fourth approach complementary to all three but particularly to bottom-up impact evaluation is using policy theory. The general principle is that a likely theory is drawn up on how the policy instrument should achieve its targeted impact in terms of energy efficiency improvement. Several terms are used for this kind of approach, including logic model analysis (Megdal 2005), realistic evaluation theory (SRC 2001), program theory (IEA 2005), theory-based approach (Rufo et al. 1999; Goldstone et al. 2000), and intervention theory (Dunn 2003). The advantage of this approach is that insight can be gained on the full implementation process, including explanatory factors behind the impact.

Applied methodology: theory-based policy evaluation

Theory-based policy evaluation

The theory-based approach is not new and has been used numerous times to evaluate policies. The method

of “theory-based” policy evaluation is extensively described and illustrated in Rossi et al. (2004). The application of the method for evaluating energy efficiency policies in Europe has been limited. The previous sections showed that there is a longer tradition with these types of evaluations in evaluating DSM programs in the USA. In Europe, the method was, e.g., applied to evaluate energy efficiency policies applied in the built environment (Joosen et al. 2004) and to perform a mid-term evaluation of the Reduction Plan on Non-CO₂ Greenhouse Gases (Harmelink et al. 2006). However, this did not result in a tradition to systematically set up guidance for the future evaluations of energy and climate change policies.

The theory-based approach requires an iterative process of program design, evaluation, and redesign based on the lessons learned from the initial program (see Fig. 1). In practice, this means that theory-based policy evaluation establishes a plausible theory on how a policy instrument (or a package of instruments) is expected to result in energy efficiency improvements, and who is expected to take action at which point in time. The basic idea is to unravel the whole policy implementation process. Through this unraveling, insight is gained on ‘where

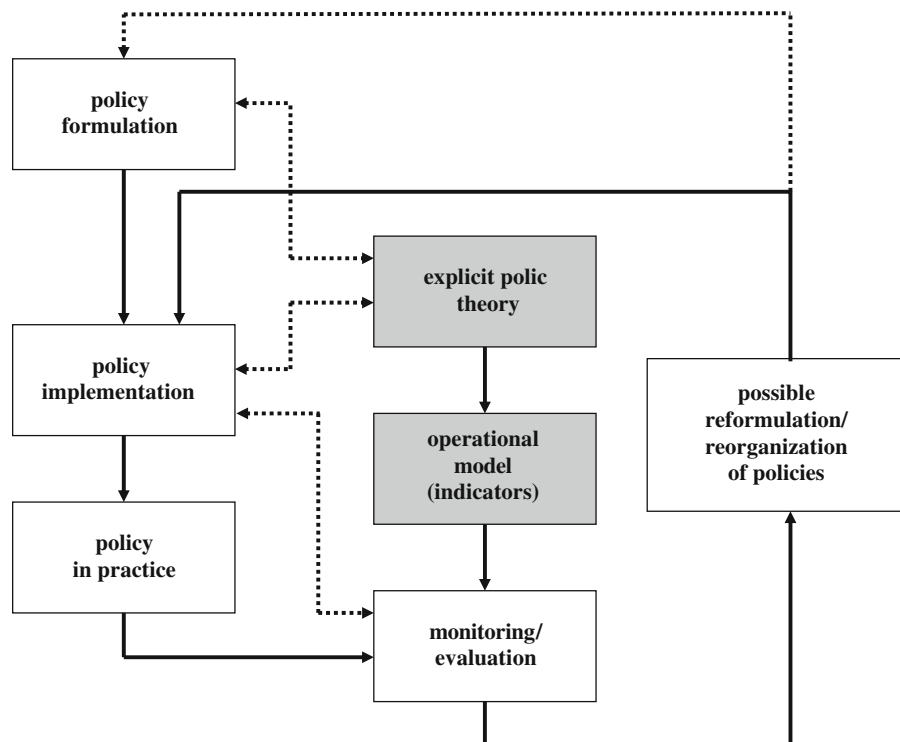
something went wrong in the process of policy design and implementation’ and ‘where the keys are for improving the impact and cost effectiveness.’

Practical framework to evaluate policy instruments

Within the framework of the AID-EE Project, the theory-based policy evaluation method was translated into a practical six-step approach that was applied to evaluate the 20 case studies. An earlier version is described in Harmelink et al. (2005). The approach applies the following steps to evaluate individual instruments or packages of instruments:

1. In the first step, the policy instrument or policy package is characterized. This includes, among others, a description of targets, the period the policy instrument was active, the target groups, the policy-implementing agents, the available budget, the available information on the initially expected energy savings impact, and the cost effectiveness of the instrument.
2. In the second step, a policy theory is drawn up. A policy theory includes all the assumptions on the

Fig. 1 Outline of the policy cycle and the role of the program theory in the policy cycle



way the policy instruments should reach its targeted impact. A policy theory can be either explicit or implicit. In the ideal case, an explicit theory is available. This means that the policy makers have clearly described how they think the policy instrument is going to work before its implementation. That is, that they have clearly stated which actor needs to take action and that they have stated the expected outcome of each action. Often, the theory is largely implicit, and such a description is lacking. In this case, the evaluator has to try to reconstruct the policy theory. Drawing up a policy theory in practice includes documenting all implicit and explicit assumptions in the policy implementation process and mapping the cause–impact relationship, including the relationship with other policy instruments.

3. In the third step, the policy theory is translated to concrete and preferably quantitative indicators. This means that, for each assumed cause–impact relation, an indicator is drawn up to “measure” whether the cause–impact relation actually took place and to “measure” whether the change (or part of the change) that took place is due to the implementation of the policy instrument (i.e., the policy instrument was the causal force). This step also includes the development of the necessary formulas to calculate the impact and cost effectiveness.
4. In the fourth step, the cause–impact relations and the indicators are visually reflected in a flowchart. An example of such a flowchart is provided in Fig. 2.
5. In the fifth step, the policy theory is verified and, if necessary, adjusted. In step 2, the policy theory was drawn up with the help of available (official) documents or experiences with similar instruments. In the fifth step, the policy theory is verified through interviews with policy makers, implementing agents, and other actors involved in the implementation and monitoring of the policy instrument.
6. In the sixth and final step, (a) available information is gathered and analyzed to draw up the indicators, (b) conclusions are drawn on the energy savings impact and cost effectiveness of the policy instrument using the formulas and indicators, (c) analyses are made on the success and failure

factors attributed to the analyzed instruments, and (d) recommendations are formulated to improve the energy savings impact and cost effectiveness.

Characteristics of the selected instruments

A great variety of policy instruments are in place in EU member states, on the EU level, and in countries outside the EU to stimulate energy efficiency improvement in different sectors. Instruments range from direct regulations, financial incentives, and voluntary agreements. Within the AID-EE Project, we aimed for a representative selection of instruments to evaluate; i.e., the selected instruments should be a good representation of the great variety of different types of policy instruments applied in the different sectors to improve energy efficiency. The following criteria were applied for selecting the 20 case studies:

1. Selected instruments should be aimed at achieving substantial energy savings and/or be aimed at market implementation of energy-efficient technologies at the national level.
2. Selected instruments should be aimed at the implementation of energy end-use efficiency improvement measures (i.e., we excluded, e.g., combined heat and power policies).
3. Selected instruments should be applied in sectors that are covered under the ES Directive.
4. Balanced breakdown of selected instruments among different sectors.
5. Balanced breakdown of selected instruments among different types of instruments; i.e., the total package of selected instruments should be a good representation of the existing variety of implemented instruments.
6. Some monitoring data should be available.

Table 1 presents the instruments that were selected as case studies in the AID-EE Project. The instruments are grouped by type of instrument: regulation, financial, informative, voluntary agreements, and procurement. It must, however, be noted that most instruments come in a package (e.g., regulation that is linked to information campaigns and financial incentives) so that it is not always easy to clearly put the instruments in a specific

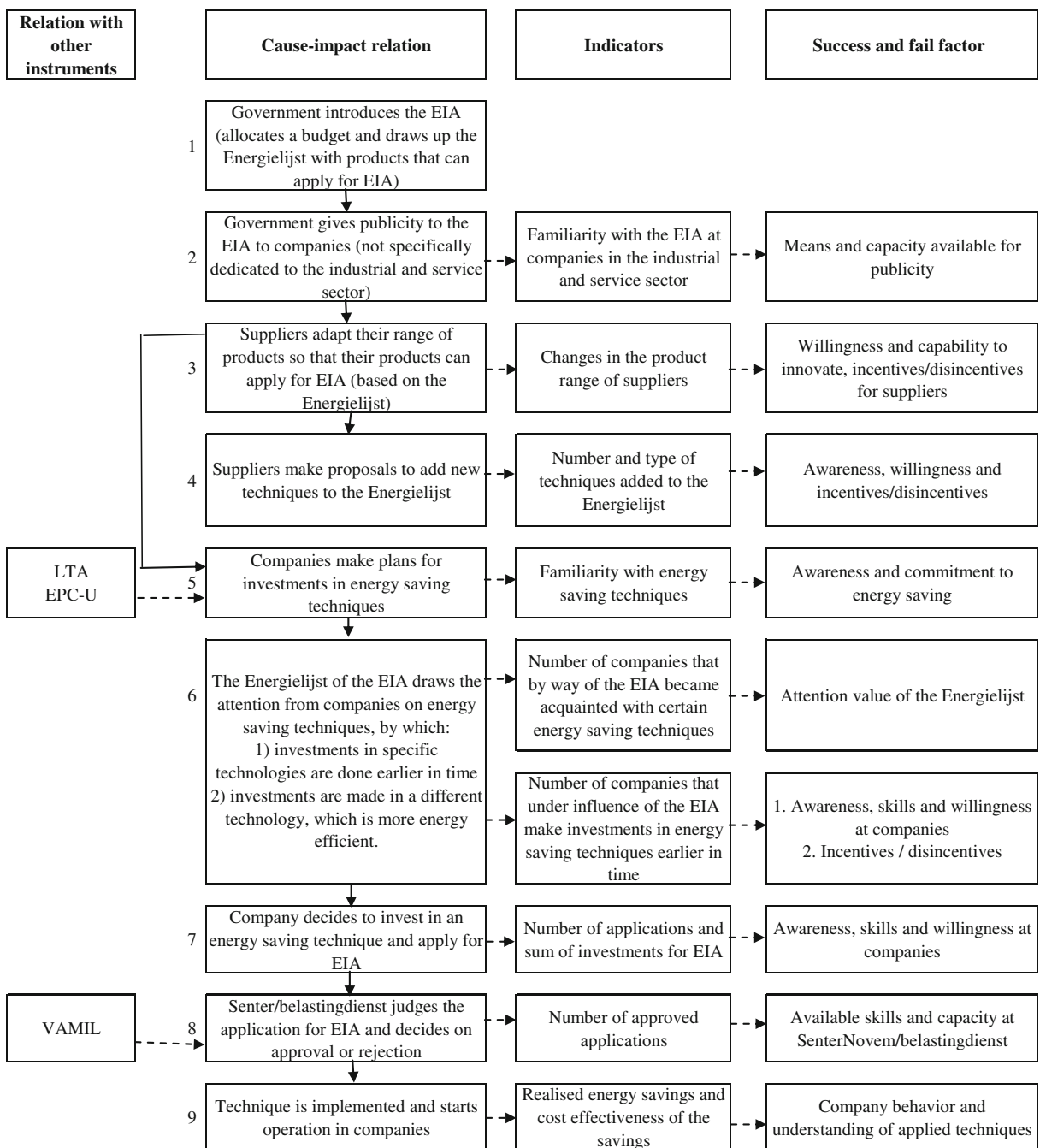


Fig. 2 Example of a flow chart for the Energy Investment Deduction Scheme (EIA) in The Netherlands showing cause–impact relations, indicators, success and failure factors, and interactions with other instruments

category. Hence, the categorization is somewhat arbitrary. This also implies that we did not only assess individual instrument but in most cases packages of instruments.

Results

This section presents the overall findings of the 20 case studies with respect to data availability, targets and target

Table 1 Overview of the analyzed instruments

	Instrument	Targeted sector	Country
Regulation	Stepwise increase of the Energy Performance Standard for new buildings Including subsidies for demonstration projects	Households and services	The Netherlands
	Energy Efficiency building regulation	Households	Italy/Carugate (province of Milan)
	Mandatory and recommended measures to improve the energy efficiency of new and renovated buildings		
	Particular focus on gas- and oil-heated buildings		
	Energy Efficiency Commitment (EEC)	Households	UK
	Obligation on gas and electricity suppliers to achieve mandatory targets for energy savings in the residential sector		
	This is a framework for large financial and information actions operated by the energy suppliers toward the target groups		
	Rational use of energy public service obligation	Households, services and industry	Belgium/Flanders
	Obligation on electricity grid companies to save energy at the end-use level (domestic and non-domestic)		
	Stimulating action (financial support) sensitizing and informing action toward target groups		
	Top runner	Households, services, transport	Japan
	Compulsory energy performance standards for a variety of products (domestic appliances, lighting, air conditioners, cars, etc.)		
	Energy labeling of domestic appliances Including a rebate	Households	The Netherlands
Obligation on having an energy manager	Services and industry	Italy	
To guarantee that companies that have an energy use above 10,000 toe/year (industrial, tertiary sector) or above 1,000 toe/year (public sector) employ an expert who deals with the analysis of energy flows, promotes energy efficiency measures, etc.			
Financial	KfW soft loan program	Mainly households	Germany
	Reduced interest rates for energy savings investments to modernize buildings		
	Energy investment deduction scheme	Services and industry	The Netherlands
Informative	Fiscal instrument, which allows companies to deduct part of their investment costs in energy efficiency equipment from the profit tax		
	Local Energy Advice Program	Households and services	Sweden
	Enable every municipality in Sweden to employ an energy adviser		
	The task of the energy adviser is to give objective advice on energy savings and renewable energy to households and local companies and organizations		
	Energy audit program	Services and industry	Finland
	Subsidies to companies and organizations to carry out energy audits for their buildings and processes		
	Closely linked to the Finnish voluntary agreement scheme		
	Industrial energy efficiency network	Industry	Norway
	Identification and realization of industrial energy savings potentials		
	Network members can obtain grants to analyze the potential for energy savings and benchmark their performance against other companies		
Energy concept for trade and industry sectors	Services and industry	Germany/North Rhine Westphalia	
Concerted development of concrete and sector-specific measures for energy savings in small-and medium-sized enterprises			
Individual advice services	Services and industry	Germany/North Rhine Westphalia	
To increase awareness and give advice to small- and medium-sized enterprises on energy saving options			
Eco-driving	Transport	The Netherlands	

Table 1 (continued)

	Instrument	Targeted sector	Country
	Information campaign on the concept of eco-driving (energy efficient driving) including training of drivers, eco-driving as part of the drivers' curriculum, in-car devices		
	Federal Energy Management Program (FEMP)	Public services in the federal sector	USA
	Variety of instruments to improve the energy efficiency of federal agencies		
	Governance by example: setting an example for other building owners and consumers, providing a market entry-point for new technologies and applying the federal buying power to expand and focus demand for energy efficiency products		
Voluntary agreement	Voluntary agreements on energy efficiency in trade and industry	Services and industry	Denmark
	Part of the Danish Green tax package (green taxes, subsidies, and VA scheme)		
	ACEA agreement	Transport	Europe
	Voluntary agreement with the European car manufacturers to make cars more efficient		
Procurement	Energy	Households	Europe
	Cooperative procurement program of energy agencies and research institutes to stimulate the market for EE domestic cold appliances		
	BELOK	Services	Sweden
	Procurement program for commercial buildings for the development of energy efficient systems and products		

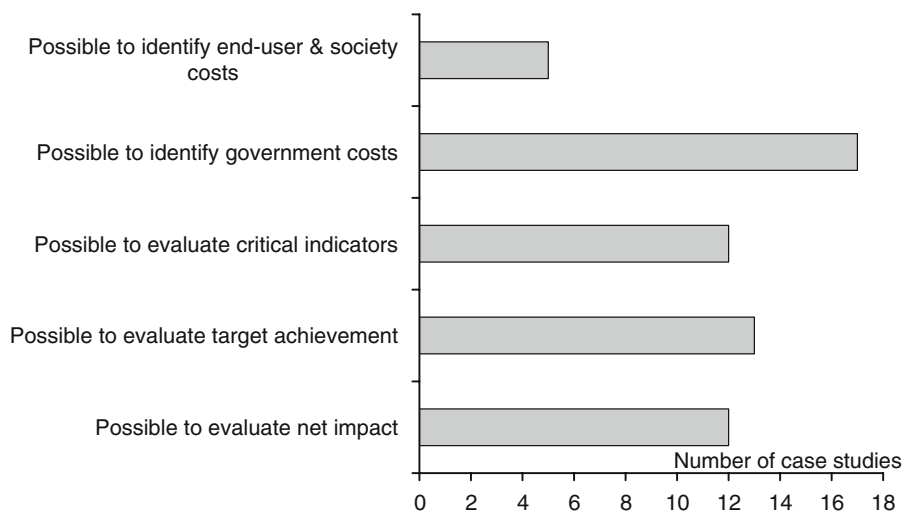
achievement, impact on energy savings, cost effectiveness, and typical success factors per instrument. Individual case studies are available at <http://www.aid-ee.org>.

Data availability

Figure 3 shows the extent to which it was possible to assess the energy savings impact, the target achieve-

ment, critical indicators determined in the policy theories developed for the instruments, the side effects, and the costs. The main observation from the case studies is that most instruments lack a comprehensive monitoring system. The availability and quality of monitoring data turned out to be much lower than expected at the start of the project. Data to assess target achievement and energy savings impact

Fig. 3 Availability of monitoring information for the 20 different case studies



were available for roughly half of the evaluated instruments. Quantitative information on indicators that could explain success or failure is hardly available. For 17 of the evaluated instruments, figures on government costs were available. Information on costs for end users and society were only available for a small number of the evaluated instruments.

Targets and target achievement

A target was defined as a specific, either quantitative or qualitative objective that has been set at the time the policy instrument was introduced. Target achievement was defined as the extent to which a policy instrument achieved its stated target(s). Figure 4 shows that, from the selected case studies, only five instruments focus on energy savings only and do not explicitly seem to have other objectives. Energy savings policies and CO₂ emission reduction policies can often not be separated. In seven case studies, CO₂ emission reduction is the primary objective, and it is mentioned as an additional objective in three case studies. Market transformation is an important additional objective in seven case studies.

Table 2 presents an overview of some of the characteristics of the evaluated policy instruments with respect to the targets set (quantitative or qualitative), the evaluated period, and the target achievement. In general, for regulatory instruments and voluntary agreements, quantified targets are set,

whereas this quantification lacks for most of the informative instruments. Six of the evaluated instruments reached their target. We also observed that, for most instruments, quantitative targets are set for the targeted end-year and generally not for intermediate years.

Impact on energy savings

The impact (in the literature also referred to as effect) of a policy instrument was defined as the extent to which a policy instrument made a difference compared to the situation without a policy instrument (business as usual). Business as usual is defined as the development of energy consumption/demand in the absence of the evaluated (package of) policy instrument(s). In Fig. 5, the energy-saving impact is presented as the annual energy efficiency improvements in the evaluated period. The impact was corrected for free-rider effects. Other gross-to-net correction factors due to rebound and spill-over effects were not taken into account due to lack of data. Total energy use of the targeted sector was taken from the European Energy and Transport Trends 2030 (EC 2003).

$$Ir = \frac{Es}{t \times Eu}$$

Ir impact of energy savings in percent per year
Es total cumulative energy savings in the end-year

Fig. 4 Type of target set for the 20 different case studies

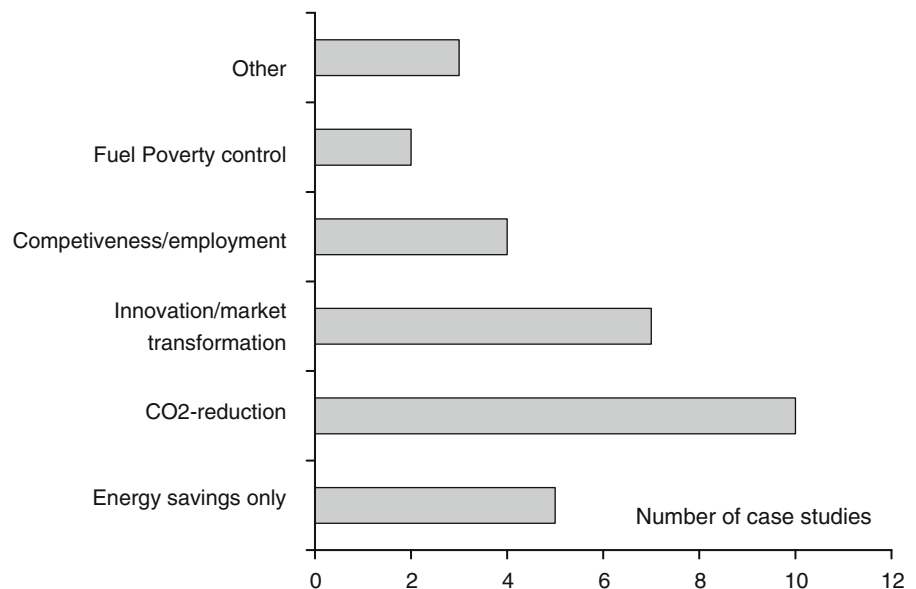


Table 2 Evaluated period, target, and target achievement for the 20 case studies

	Instrument	Evaluated period	Target	Target achieved
Regulation	Energy performance standard for buildings (NLD)	1996–2004	2	A
	Building regulation (ITA)	2003–2005	2	D
	Energy Efficiency Commitment (UK)	2002–2005	2	A
	Mandatory targets on energy consumption (BEL)	2003–2004	2	A
	Top runner (JAP)	1999–2005	2	A
	Labeling of domestic appliances (NLD; rebate)	1995–2004	2	C
	Obligation on having an energy manager (ITA)	1999–2003	1	D
Financial	Soft loans for building modernization (GER)	1996–2004	2	B
	Energy investment deduction scheme (NLD)	1997–2004	1	D
Informative	Local energy advice (SWE)	1998–2004	1	D
	Energy audits program (FIN; subsidy), public services	1992–2004	2	A
	Energy audits program (FIN; subsidy), private services	1992–2004	2	B
	Energy audits program (FIN; subsidy), industry services	1992–2004	2	C
	Industrial energy efficiency network (NOR)	1996–2004	1	D
	Energy concept for industry sectors (GER)	1996–2003	1	D
	Individual advice services (GER)	1990–2005	1	D
	Eco-driving (NLD)	2000–2004	2	B
	FEMP (USA)	1985–2004	2	C
	VA	Voluntary agreements on energy efficiency (DEN; subsidies)	1996–2003	2
ACEA covenant (EUR)		1998–2003	2	C
Procurement	Energy (EUR)	1999–2004	1	D
	BELOK (SWE)	2001–2005	1	A

1 qualitative target or no target exists for this instrument; *2* quantitative target exists for this instrument; *A* target for this instrument was achieved or overachieved; *B* target has not been achieved; *C* target year has not been reached yet, unclear whether target achievement is on track; *D* due to a lack of a quantified target, target achievement cannot be assessed

of evaluations *corrected for free-rider effects* in PJ

t evaluated period in years

Eu total energy use of the targeted sector end in end-year of evaluation in PJ

Table 2 shows the evaluated period, target, and target achievement for the 20 case studies.

Figure 5 shows the calculated annual energy efficiency improvement with a breakdown by different instruments. The transport sector shows the lowest improvement rates. It can also be observed that there are almost no instruments in place that exclusively focus on the service sector. All instruments that target the service sector also address either industry or households. Instruments that exclusively focus on the service sector are rare. As the Flemish and many US commercial sector programs show, it is not more difficult to achieve savings in the service sector than elsewhere. Figure 5 does not show clear differences in savings for the different types of instrument. One should note that, for some instruments, such as the

Danish VA scheme, the Dutch labeling, and the Finnish audit program, the results are given for the policy package and not for individual instruments. This might explain why the energy audit and the labeling scheme (both including financial incentives) have a higher improvement rate than the other mainly informative instruments. Furthermore, in some sectors and countries, there was relatively much “low-hanging fruit,” i.e., highly cost-effective measures that are not implemented due to various barriers. In such cases, high-energy efficiency improvement rates may be possible. Some instruments focus on just a small part of the sector, which makes the energy efficiency improvement rate at sector level rather small. An example is an energy performance standard for new buildings that only slowly penetrates because the demolition rate in most countries is rather low.

It must be noted that data on the energy-saving impact of instruments are surrounded by relatively high uncertainties because of limited availability of data. This is reflected with the error bars in Fig. 5. Due to lack of monitoring data, several assumptions

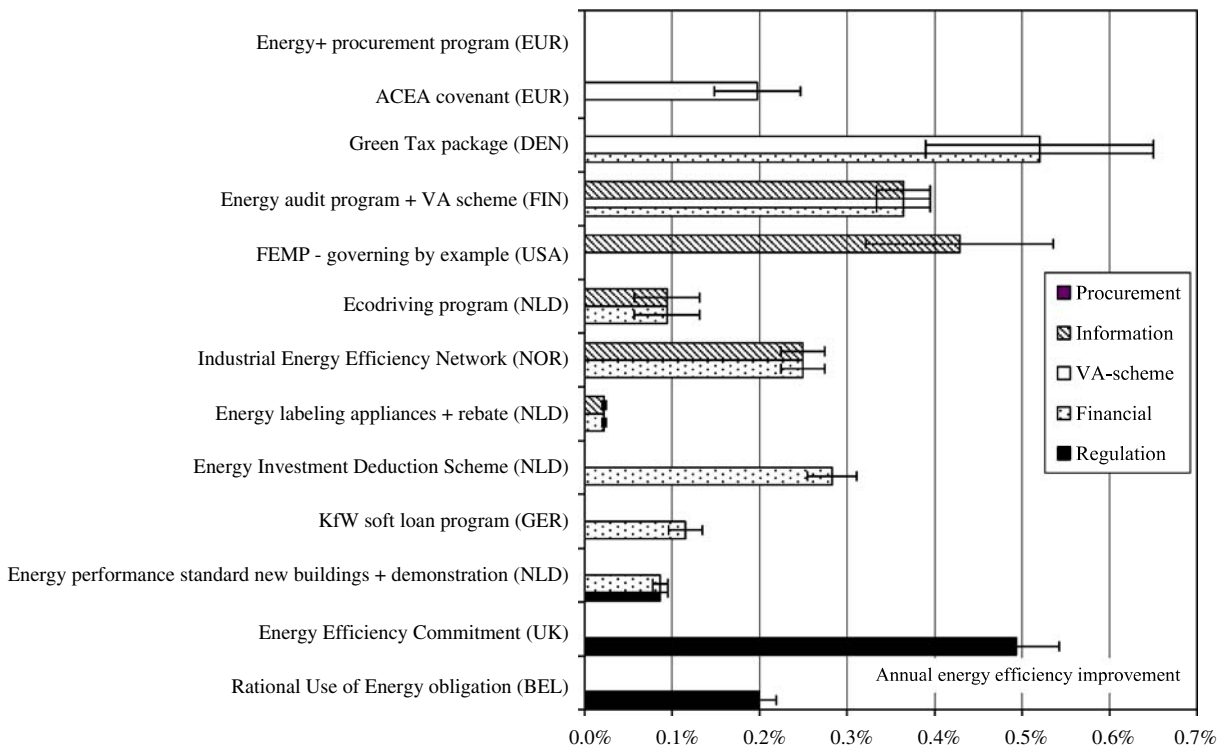


Fig. 5 Annual percentage energy savings improvements per instrument in the evaluated period

had to be made on “real-life” performance of energy-saving measures (actual energy efficiency performance, operational hours, use of buildings and appliances, etc.).

Figure 5 shows the annual percentage energy savings per instrument and sector in the evaluated period (see Table 2 for the evaluated period). Error bars reflect uncertainty in the calculated impact.

Most evaluated instruments are part of a policy package (see Table 1). Often links between instruments are intended and meant to increase the impact of the whole package with respect to energy savings. In general, it is difficult to determine the isolated impact of a single instrument in a policy package. Informative instruments, which are generally implemented to support other instruments, constitute a good example. Their isolated impact is generally small or even zero. However, our analysis shows that both regulatory and financial instruments, as well as voluntary agreements, would not be so effective without informing target groups on their obligations, financial benefits, etc. Literature on the reinforcing or mitigating effect of policy instruments is scarce, and quantitative results are mostly lacking. Boonekamp

(2005) developed a method to analyze the interaction between two instruments. This method was further elaborated by Michelsen (2005) for policy instruments providing incentives for the procurement of energy-efficient cold appliances and complemented with empirical evidence by interviewing experts.

Cost effectiveness

Cost effectiveness refers to the ratio between the energy-saving impact and the amount of money needed to achieve this impact and can, e.g., be expressed in euros per GJ or kWh saved. For a complete assessment of the cost effectiveness of a policy instrument, costs and benefits should be assessed from different perspectives:

- End user. The costs for the end user provide an indication of the costs as experienced by the end user responsible for the implementation of the energy efficiency measure. These costs are defined as all additional costs that have to be made by the end user compared to the reference situation in case the evaluated energy efficiency policy instrument

would not have been in place. The benefits include all savings on the energy bill.

- Society. The costs for society provide an indication of the costs as experienced by the society as a whole for the implementation of the energy efficiency measure. The main difference between cost–benefit analysis from an end-user perspective and from a social perspective is the time preference. The social perspective is translated into a discount rate that is generally much lower than cutoff discount rates used by end users. In the case of cost–benefit analysis from the social perspective, the discount rate is called the social discount rate.
- Governments. Costs for the government are defined as all expenditures that have been made by the government, which can be related to the implementation of the evaluated energy efficiency policies. Government expenditure includes budgets for subsidies, grants for research and development, costs for monitoring, and the administrative machinery. However, this also includes reduced government income due to fiscal measures and lowered energy tax income.

In the case studies, we have tried to determine the cost effectiveness for the three perspectives. However, due to the limited availability of data, it was impossible to make a useful comparison of the cost effectiveness from the perspective of the end user and the society. Government expenditures were typically between 0 and 5 EUR/GJ saved in the cases where this number could be determined. Although cost efficiency is an important evaluation criterion, it was beyond the means of the case studies to collect primary data for assessing this. In some cases, evaluations that included cost-efficiency assessments existed, but due to lack of transparency and documentation, and apparent differences between approaches, a comparison between estimates is not meaningful.

Success and fail factors

Using a policy theory approach facilitates a systematic and structured assessment of the cause–impact relationships in the implementation of policy instruments and, hence, the identification of success and fail factors. The 20 policy instruments evaluated in the AID-EE Project are of different types; they target different sectors, and they are implemented in different organizational and

institutional contexts. Nevertheless, it is possible to make some general observations based on the recurrence of certain observations made independently in the different case studies. Table 3 in Appendix indicates in what circumstances that different types of instruments are suitable and, derived from the case studies, lists a number of specific characteristics that have been identified as potentially important determinants of success or failure.

There are also a few general factors in the process of design and implementation of policy instruments that appear as important. These interrelated factors include the existence of clear goals and a mandate for the implementing organization, the ability to balance and combine flexibility and continuity, the involvement of stakeholders, and the ability to adapt to and integrate adjacent policies or develop consistent policy packages.

It is not possible to conclude from the case studies whether it is important that the implementing agency is independent or not from the regulator. However, there is a strong indication that an organization or program with a clear mandate, responsibility, and adequate resources is a first and foremost prerequisite for success. The continuity of a program is important and can be ensured either by strong political backing within central bureaucracies or through creating more independent implementing agencies. Continuity in terms of stable and predictable conditions has been identified as an important success factor in several case studies. However, flexibility in the sense of being able to adapt to changing conditions or to remove potential fail factors in the implementation process is equally important. Technical and market developments, organizational changes, and changes in other policy areas may motivate adjustments. In several cases, the ability of the implementing agency to quickly remove specific barriers was noted as important. Such barriers include lack of information, tools, and skills that require the development of guidelines, procedures, analytical tools, education, etc.

The involvement of stakeholders in the design and implementation of policy instruments is noted as an important success factor in several cases. It serves the dual purpose of increasing the acceptance and, at the same time, improving efficiency by ensuring that stakeholder needs and preferences are considered throughout the process. Stakeholders include both the targeted actors and the various other actors that may be affected, e.g., energy consultants, equipment suppliers,

and third-party auditing companies. A common trait is that the stakeholders appreciate simple and clear rules, as well as short lead times in cases where decisions or approvals have to be made by the implementing agency. Stakeholder involvement facilitates flexibility as noted above but also makes it easier for the implementing agency to benefit from other developments and conditions, for example, the existence of environmental management systems. An implementing agency with the ability to communicate well with all stakeholders but also with the mandate to adjust and evolve the policy instrument stands a better chance of being successful.

Conclusions and discussion

The case studies and the overall assessment of results presented in this paper lead to a number of observations concerning the usefulness of the policy theory approach, the experience with energy efficiency policy instruments, as well as the general success and fail factors. Based on these observations, policy implications in the context of the ES Directive are discussed.

The evaluations of 20 energy efficiency policy instruments applied across Europe, the USA, and Japan show that:

- Energy efficiency policies often lack quantitative targets and clear timeframes.
- Policy instruments often have multiple and/or unclear objectives.
- The need for monitoring information often does not have priority in the design phase.
- For most instruments, monitoring information is collected on a regular basis. However, this information is often insufficient to determine the impact on energy saving, cost effectiveness, and target achievement of an instrument.
- Monitoring and verification of actual energy savings have a relatively low priority for most of the analyzed instruments.
- There is no such thing as the ‘best’ policy instrument. However, typical circumstances in which to apply different types of instruments and generic characteristics that determine success or failure can be identified.

Despite the weaknesses noted above, as well as the diversity of instruments and contexts in which they were used, it is possible to make a few general

observations concerning success factors in the process of design and implementation of policy instruments. Success factors include:

- The existence of clear goals and a mandate for the implementing organization.
- The ability to balance and combine flexibility and continuity.
- The involvement of stakeholders.
- The ability to adapt to and integrate adjacent evolving or new policies or develop consistent policy packages.

The instruments were evaluated by applying a practical framework based on theory-based policy evaluation. With the development of this framework, we aimed to contribute to a further harmonization of evaluation processes and create comparable evaluation outcomes among the case studies. We conclude that the theory-based policy evaluation method has several benefits over other ex post evaluation methods since:

- The whole policy implementation process is evaluated, and the focus is not just on the final impacts (i.e., realized energy savings).
- Through the development of indicators for each step in the implementation process, the “successes and failures” can be determined to the greatest extent possible.
- By applying this approach, we not only learn whether policies are successful or not but also why they succeeded or failed and how they can be improved.

We experienced a number of practical problems that in real cases often make it difficult to exactly follow all steps of the developed methodology. These problems are mainly associated with the fact that evaluation and monitoring aspects did not have priority in the design and implementation phase of our case studies. Such problems include:

- Lack of monitoring data and information. Baseline data or data on penetration rates, savings, awareness, etc., are often difficult and sometimes impossible to reconstruct.
- Lack of time and resources mean that compromises have to be made regarding data collection and reconstruction. Ideally, decisions should be made early in the policy process regarding level of ambition and focus of the evaluation to identify important indicators and data.

- The policy theory is not clear and not explicitly formulated during the design and implementation phase. When it has to be reconstructed, there is a risk of misinterpretation unless it is carefully checked with several informants.
- In many cases, it is difficult to determine clear cause–effect relations. Integrating policy theory in the policy design phase would at least ameliorate some of these difficulties.
- It may be difficult to identify success and fail factors, and determine their importance. Different respondents sometimes have different views on this, but through an iterative process including several respondents as well as factual information, a reasonably accurate representation can be constructed.

Based on the analysis and results, it is worth making a few observations in the context of the ES Directive and the 20% target for 2020. The ES Directive requires that savings are monitored and verified. Our analysis shows that experiences in this area from existing policy instruments in Europe are limited. Hence, the development of harmonized evaluation methods under the ES Directive starts from a relatively low level. It is also in the initial stage entirely focused on quantifying the savings. The analysis presented here shows that the policy theory approach can add considerable value in the process of policy design, implementation, and evaluation. Mainly, this results from systematizing the search for cause–

impact relationships and explanatory factors behind the impact of instruments. In the design phase, it helps policy makers think about the whole implementation process, the relation and overlap with other policy instruments, and the relevant indicators to monitor so that evaluation becomes an integral part of the policy instrument and not a later add-on. During the implementation process, the approach can help identify problems and barriers at an early stage and allow the policy instrument to evolve and adjust to changing technical, organizational, economic, and other conditions. For verifying the savings, a policy theory approach can improve the quality of evaluations by ensuring that data needs and data gathering is an integral part of the policy. This is important not least for handling gross-to-net correction factors. Effective and efficient energy efficiency policies in terms of delivering large savings at low cost will be crucial for reaching energy and climate policy goals for 2020 and thereafter. For this purpose, evaluations must go beyond determining merely the savings as stipulated in the ES Directive and include approaches that facilitate continuous policy learning and policy transfer.

Acknowledgment The authors would like to thank the other project partners of the AID-EE Project, the Wuppertal Institute for Climate, Environment and Energy, and the Politecnico de Milano for their participation and contributions to the results of the AID-EE Project. Without their efforts, we would not have been able to compile this paper. We also thank two anonymous reviewers for valuable comments.

Appendix

Table 3 Typical circumstances in which to apply different types of instruments and characteristics that determine the success of the instrument

Type of instrument	Typical circumstances in which to apply this instrument	Characteristics that typically determine the success
Energy performance standards for buildings, cars, or appliances	When dealing with a target group which is: unwilling to act (e.g., voluntary agreement of producers not fulfilled) and difficult to address (e.g., landlord–tenant problem)	Is the standard well justified? For example, through life-cycle cost studies Is the target group well prepared for the standard? For example, through information campaigns, demonstration projects, feasibility studies, training programs, etc. Is the target group sufficiently skilled to apply the standard?
	When aiming at removing the worst products or services from the market with regard to energy consumption	Is there resistance among the target group to apply the standard? Are there sufficient resources (knowledge, capacity, time, budget, priority) in place to enforce the legislation?

Table 3 (continued)

Type of instrument	Typical circumstances in which to apply this instrument	Characteristics that typically determine the success
Mandatory targets/tradable certificates for (demand-side) energy savings for energy companies	When aiming at energy savings in the households or services sector, i.e., large target groups being difficult to address by energy efficiency services When knowledge, financial, and institutional barriers play a role As an alternative or complement to an energy saving fund	Are there penalties in place for non-compliance?
		Are the penalties at a sufficiently high level to stimulate meeting the standard?
		Is the standard timely adjusted to technology progress?
		Is the target clearly set beyond business-as-usual?
Labeling of appliances, cars, buildings	When there is a knowledge/information barrier When dealing with large consumer or service sector groups When dealing with rather uniform technologies When there are large differences in energy performance between similar units	Is measurement and verification of savings possible at low cost, e.g., by standardization of energy-saving measures?
		Is the cost-recovery mechanism (energy companies' costs passed to end users) clear and transparent?
		Are there penalties in case of non-compliance (or are there other incentives in place to prevent non-compliance)?
		Are penalties set at such a level that target achievement is stimulated?
		Are financial incentives needed to stimulate households and companies to implement EE measures
Financial / fiscal instruments such as soft loans, subsidy schemes, investment deduction schemes, rebates	When there is a financial barrier in place When an informative instrument (e.g., energy audit) needs financial incentives to attract the target group	Is the market for tradable certificates transparent and reliable?
		Is there undesired overlap with other instruments?
		Is it planned to adjust the label to technology progress and market transformation?
		Is the label well justified by respective life-cycle cost studies?
		Is the target group timely and sufficiently informed? For example, through information campaigns
		Is the label clear and transparent?
		Are there complementary incentives (eco-tax, subsidy, tax exemptions) for stimulating action?
		Is the target group aware of the existence of the instrument?
		Is the financial support sufficient to attract new investments or to carry out energy audits?
		Is the annual budget for the instrument linked to the target?
		Is the procedure for getting financial support sufficiently known by the target group and simple enough?

Table 3 (continued)

Type of instrument	Typical circumstances in which to apply this instrument	Characteristics that typically determine the success
Energy tax / energy tax exemption	When dealing with large target groups	Is it clear for the target group which technologies are eligible for financial support? Is the list of technologies regularly updated to limit free riders? Is the instrument implemented for a long time period to ensure security for investors? Is the target group well informed on existence and planned future development of the energy tax?
	When aiming to internalize external costs	Is use of tax income properly justified and marketed to market actors? To what extent does the energy tax take account of global or European-wide competition aspects (e.g., by tax exemptions for large industries)? To what extent are energy tax exemptions used as an incentive for implementing EE measures (e.g. in a voluntary agreement scheme)
Information / knowledge transfer / education / training	When there is a knowledge barrier	Is the information well linked to the customer type within the target group?
	When dealing with large target groups	Is the information clearly linked to other instruments (regulation, financial/fiscal, voluntary agreement, etc.)?
Energy audits	When there is a knowledge barrier for buildings and production facilities	Is the target group well-informed about existence of instrument? Is the target group well-informed about benefits and costs of instrument and of energy-saving measures identified (e.g., through demonstration projects)? Is the energy audit targeting all relevant energy end uses? Is the energy audit producing an estimate of energy cost savings and investments for the recommended measures?
		Is the energy audit scheme linked to financial incentive, soft loan, voluntary agreement, and/or energy performance contracting schemes?
Voluntary agreements to save energy (industry, services sector) or improve energy efficiency (e.g. cars or appliances)	When dealing with a small number of actors with which you need to negotiate or a strongly organized sector	Is the target group motivated to participate in the voluntary agreement?
	When there is much relatively cheap saving potential (low hanging fruit)	Is the target set beyond business-as-usual? Are there penalties in case of non-compliance (or are there other incentives in place to prevent non-compliance, e.g. a rebate on energy tax, or is there a regulatory threat in case of non-compliance)? Is there a good monitoring system in place?

Table 3 (continued)

Type of instrument	Typical circumstances in which to apply this instrument	Characteristics that typically determine the success
Co-operative procurement programme	<p>When there are sufficient possibilities to bundle large buyers of EE technologies</p> <p>When there is a limited number of market actors supplying EE technologies</p> <p>When potentials for further development and market transformation of new technologies are large enough</p>	<p>Are supporting instruments in place (such as audits, energy monitoring systems, demonstration projects, financial incentives)?</p> <p>Is the program management qualified and engaged? Can the buyers and suppliers group be motivated in principle? Is the buyers group involved in the program set up?</p> <p>Is the buyers group sufficiently sized? Are the results of the program well documented to facilitate market deployment? Is the program well tuned with other policies (energy efficiency standards, labeling, research, and development)?</p>

References

- Blumstein, C., Goldstone, S., & Lutzenhiser, L. (2000). A theory-based approach to market transformation. *Energy Policy*, 28(2000), 137–144.
- Bongardt, D., & Kebeck, K. (2007). *New governance or symbolic policy?—Evaluation and recommendations for the agreement between the European Commission and the Automobile Industry*. Paper presented at the ECEEE Summer Study 2007, France, June.
- Boonekamp, P. (2005). *Improved methods to evaluate realised energy efficiency*. PhD thesis, The Netherlands, University of Utrecht.
- Dunn, W. (2003). *Public policy analysis an introductory* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
- EC (2003). *European energy and transport: Trends to 2030*. Brussels: European Commission.
- EC (2006a). *Action plan for energy efficiency: Realising the potential*. Brussels: European Commission 19.10.2006 COM(2006)545 final.
- EC (2006b) *Directive 2006/32/EC of the European Parliament and the Council of 5 April 2006 on Energy End-use Efficiency and Energy Services and Repealing Council Directive 93/76/EEC*.
- EC (2007). *An energy policy for Europe*. Brussels: European Commission 10.1.2007 COM(2007)1.
- Ecofys, Lund University, Politecnico di Milano, Wuppertal Institute (2007). *Success and failure in energy efficiency policies. Ex-post evaluation of 20 instruments to improve energy efficiency across Europe*. Reports published within the framework of the AID-EE Project by Ecofys, Lund University, Politecnico di Milano, Wuppertal Institute (<http://www.aid-ee.org>).
- Goldstone, S., Rufo, M., & Wilson, J. (2000). *Applying a theory-based approach to California's non-residential standard performance contract program*. Lessons learned by Goldstone and John Wilson, California Energy Commission; and Michael Rufo, Xenergy Inc. from the 2000 ACEEE Summer Study on Energy Efficiency in Buildings.
- Hall, N., Roth, J., & Best, C. (2006). *California energy efficiency evaluation protocols: Technical, methodological, and reporting requirements for evaluation professionals*. Prepared for the California Public Utilities Commission by The TecMarket Works Team, Oregon, USA, April 2006.
- Harmelink, M., Joosen, S., & Blok, K. (2005). *The theory-based policy evaluation method applied to the ex-post evaluation of climate change policies in the built environment in The Netherlands*. Paper presented at the ECEEE Summer Study 2005, May.
- Harmelink, M., Joosen, S., Eising, K., & de Visser, E. (2006). *Mid-term evaluation of the programme on non-CO₂ greenhouse gases*. The Netherlands: Ecofys.
- Harmsen, R., van den Hoed, R., & Harmelink, M. (2007) *Improving the energy efficiency of private road transport: the case of Ecodriving in the Netherlands*. Paper presented at the ECEEE Summer Study 2007, France, June.
- IEA (2005). *Evaluating energy efficiency policy measures & DSM Programmes. Volume I: Evaluation guidebook*. International Energy Agency Implementing Agreement on Demand-Side Management Technologies and Programmes.
- Joosen, S., & Harmelink, M. (2006). *Guidelines for the ex-post evaluation of 20 energy efficiency instruments applied across Europe*. Report published within the framework of the AID-EE Project by Ecofys, Lund, Politecnico, Wuppertal. Retrieved from <http://www.aid-ee.org>.
- Joosen, S., Harmelink, M., & Blok, K. (2004). *Evaluation of climate change policies in the built environment in The Netherlands for the period 1995–2002*. The Netherlands: Ecofys.
- Khan, J., Nordqvist, J. (2007). *Success and failure in the promotion of an increased energy efficiency in industry*. Paper presented at the ECEEE Summer Study 2007, France.
- Megdal, L. (2005). *Using program logic model analysis to evaluate & better deliver what works*. Paper presented at the ECEEE Summer Study 2005, May.

- Michelsen, C. (2005). *Exploring policy space: Interactions between policy instruments on household energy efficiency. The case of domestic cold appliances in Germany*. Sweden: Lund, IIIIEE (September 2005, Reports 2005:04).
- Nordqvist J. (2007). *The top-runner policy concept: pass it down*. Paper presented at the ECEEE Summer Study 2007, France.
- NYSERDA (2008). Retrieved February 13, 2008 from <http://www.nyserda.org/default.asp>.
- Reed, J. H., Jordan, G., & Vine, E. (2007). *Impact evaluation framework for technology deployment programs. Prepared for the US DOE by Innovol Sandia National Laboratories ogie LLC*. Lawrence Berkeley National Laboratory: Washington, USA.
- Rossi, P., Lipsey, M., & Freeman, H. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA, USA: Sage.
- Rufo, M., Prah, R., & Landry, P. (1999) *Evaluation of the 1998 California non-residential standard performance contracting program: A theory-driven approach, evaluation in transition: Working in a competitive energy industry environment*. Proceedings of the 1999 Energy Evaluation Conference, Denver.
- SRC (2001). *A European ex-post evaluation guidebook for DSM and EE Service Programmes*. Denmark: SRC International.