International Energy Agency
Implementing Agreement on
Demand-Side Management
Technologies and Programmes

EVALUATING ENERGY EFFICIENCY POLICY MEASURES & DSM PROGRAMMES

VOLUME I
EVALUATION GUIDEBOOK

BASED ON NATIONAL CASE STUDIES & NATIONAL AND INTERNATIONAL EXPERIENCES

Prepared by

Harry Vreuls
Operating Agent
SenterNovem
The Netherlands

Financed by Sweden, the Netherlands, Korea, Italy, France, Denmark, Canada and Belgium

The following countries supported this project:

Belgium
Canada
Denmark
France
Italy
Republic of Korea
The Netherlands
Sweden

Prepared by the Operating Agent, Harry Vreuls

You can contact the Operating Agent

- by post: PO Box 17, NL 6130 AA Sittard, The Netherlands
- by phone: + 31-46-4202258 and by fax +31-46-4528260
- by e-mail: h.vreuls@senternovem.nl

With contributions and input for Volume II, case examples by
Wim de Grootte (Belgium)
Peter Bach, Richard Schalburg, Kirsten Dyhr-Mikkelsen (Denmark)
Didier Bosseboeuf (France)
Ornella Celi (Italy)
Jong-Duck Kim (Republic of Korea)
Lena Neij (Sweden)
and with assistance by Mitchel Roosenburg, Kema-Exenery

This report is available from: http://dsm.iea.org

October 2005

Table of contents

SUMMARY

Many governments around the world, including those of most European countries, have signed the United Nations Framework Convention on Climate Change (UNFCCC) and the Kyoto Protocol to this Convention that came into force early in 2005. Countries have to report on their progress and the impacts of their policy measures. An EU Directive for Buildings Performance (EDBP) and the EU Directive on Energy End-use and Energy Services is due to be accepted by the end of 2005, in which the European Commission sets a target for average annual energy savings by the Member States.

Early government and utility companies' energy efficiency programmes made heavy use of economic incentives such as rebates, investment subsidies, and free services to encourage end-use customers to select energy-efficient equipment and construction designs. Evaluation of these programmes focused on the characterisation of participant response, measurement of energy savings among participants, and attribution of their actions to the programme, among other potential influences. Over the past 15 years the variety of policy measures employed by energy efficiency programme sponsors has vastly proliferated. Many energy efficiency programmes now operate on an international scale. The European Energy Label Programme, which applies to domestic appliances, operates in all EU Member States. Most Member States have also implemented the Green Lights Programme to promote the use of efficient lighting in commercial buildings. The ENERGY STAR labelling programme for office equipment, originally developed by the United States Environmental Protection Agency (US-EPA), has also been adopted by many countries across Europe. Finally, the Kyoto Protocol itself specifies three international programmes: Joint Implementation (JI), the Clean Development Mechanism (CDM), and Emissions Trading (ET).

The proliferation of energy efficiency programme approaches and organisational settings has been accompanied by increased variation in the evaluation objectives. Until fairly recently, US evaluations in particular were often oriented towards developing an estimate of energy savings attributable to a programme that was defensible in regulatory proceedings. In our opinion a useful framework for planning evaluations must relate methods not only to the nature of the programme, but also to the sponsor's evaluation objectives, the need for precision and accuracy, and budget. These last factors can be referred to as the sponsor's *level of ambition* for the evaluation. This factor is included as one of our key elements for evaluation.

In the light of these developments the authors have set the following two objectives for this *Guidebook*. Firstly, to provide guidance for the evaluation of a broad range of energy efficiency programmes currently offered by governments, energy companies, and other sponsors, with specific approaches for each major type of programme or policy measure. Secondly, to focus on providing guidance in matching research questions and methodological approaches on the one hand, to programme type and level of ambition on the other.

The *Guidebook* consists of two volumes. This Volume (I) deals with evaluation theory and recommends how evaluations for five types of policy measures and programmes should be conducted. This new approach involves organising evaluations into seven key analytic elements. Volume II covers the evaluation tradition in the various countries and a number of selected case examples on evaluations, and also provides readers with additional background information concerning the choices made, which could help them find solutions for missing elements in the theory.

Chapter 1 lays out a basic framework of selecting methods for evaluating specific energy efficiency policy measure or DSM programmes, and includes:

- A taxonomy of energy efficiency policy measures that is useful for quickly identifying the relevant evaluation issues associated with a given programme.
- Seven key analytic elements that need to be addressed in virtually all energy efficiency programme evaluations.

Chapters 2 through 6 apply this framework in providing guidance to the types of policy measures identified in the taxonomy, as well as to programmes that combine several different policy measures. Each chapter is structured according to the *seven key analytic elements:*

1. Policy measure theory used.
2. Specification of indicators for the success of a measure.
3. The baselines for the selected indicators.
4. Assessment of outputs and outcomes.
5. Assessment of energy savings and emissions reductions and other relevant impacts.
6. The calculation of costs, cost-efficiency and cost-effectiveness.
7. The level of evaluation effort.

These seven key elements are placed in a preferred evaluation framework (see figure on next page) for a normative evaluation in that inputs, outputs and impacts are organised in combination with the role an evaluation should play within the public environment.

The policy measure theory statement provides the basic framework for the evaluation. Ideally, a policy measure theory is selected (explicit or implicit) during the development of the policy measure, and information on this key element should be available to the evaluators. For the evaluation, the theory statement should also include a specification of the policy measure domain and a statement of policy measure effects hypotheses.

For the seventh key element we have devised a scale of evaluation effort or ‘ambition’ that is related to the motivations of evaluation sponsors and the rigour of the impact evaluation methods deployed. The following three levels of evaluation effort are used in discussing method selection with regard to specific types of policy measures. In practical terms, three levels of effort are defined as follows.

Level A: Comprehensive evaluation:

- outcome indicators including net behavioural change,
- impact indicators on energy savings,
- additional internal and external information sources are needed.

Level B: Targeted evaluation:

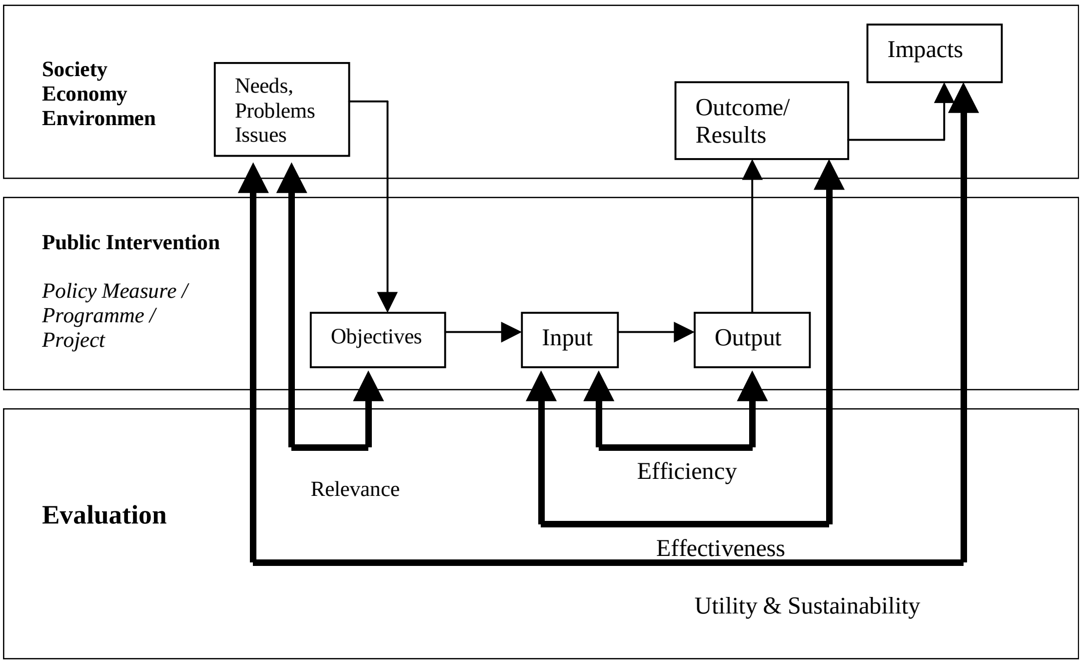
- including outcome indicators as gross behavioural change,
- some additional information sources.

Level C: Programme review evaluation:

- focus on input and output indicators,
- only using existing (written) information sources.

Chapter 2 deals with the evaluation of *regulation* policy and measures. ‘Regulation’ refers to the laws and implementing rules regarding requirements for devices to advance energy-efficient design and construction. We describe the seven key analytic elements for two categories: building codes and minimum energy performance standards.

Chapter 3 discusses the evaluation of *information* programmes. Most of the case examples in Volume II fall into this category. We have reorganised the seven subcategories based on the audience targeting and the complexity of the message or information to produce a smaller number. We group general information, labelling and information centres together and present examples for one of these in the various sections. The other group that is described in detail covers energy audits, education and training.



Evaluation Questions

- *Relevance*: To what extent are the objectives justified in relation to needs?
- *Effectiveness*: To what extents have the expected objectives been achieved?
- *Efficiency:* Have the objectives been achieved at lowest cost?
- *Utility & Sustainability*: Do the expected or unexpected effects contribute to a net increase in social welfare and sustainability?

Chapter 4 deals with *economic incentives*. This includes the broad range of subsidies, rebates, taxation, grants, loans etc. We hold the opinion that the policies dealing with bulk purchasing, technology procurement and certificate trading systems should also be included in this category. Based on the criteria ‘simple price reduction and combination of price with market developments’, we combine the various economic incentives into five groups. Three are described in detail: price reduction, taxation systems and financial arrangements, and policy and measures ensuring a minimum market.

Chapter 5 deals with the evaluation of Voluntary Agreements, and discusses the key elements for evaluating the so-called strong compliance agreements. This chapter also refers to the following chapter, as Voluntary Agreements often explicitly use other instruments such as energy audits and subsidies.

Many policy measures are not implemented as isolated measures, but as part of a mix or package of measures in order to increase the desired effect. Generally, information will always be used as part of a package because people need to be aware of what is expected of them and have the proper knowledge before they can act accordingly. Economic incentives are easily mixed with any of the other measures, while regulation and Voluntary Agreements tend to exclude each other. Chapter 6 addresses three types of policy measure packages:

1. Regulation, information, and economic incentives.
2. Voluntary Agreements, information, and economic incentives.
3. Market Transformation: economic incentives (technology procurement), information and Voluntary Agreements.

Chapter 7 lists the authors' conclusions.

Volume II starts with an overview of case examples on evaluations. These examples are included in the country reports. The individual country expert describes the national system of energy efficiency policy measures, the system for evaluating, monitoring and data collection on energy policy measures, relevant scenarios and methods for evaluating energy efficiency programmes (1995 onwards). The following case examples are included in Volume II:

Policy type	Case examples	Country
Regulation	Building codes	Belgium
	Energy Efficiency Regulations for Residential Equipment	Canada
	Energy management scheme for large buildings	Denmark
	Minimum energy performance standards	Korea
	Energy Performance Standard (EPS) for houses	Netherlands
Information	Local energy efficiency information centres	Belgium
	Energuide for houses	Canada
	Energy labelling of small buildings	Denmark
	Free-of-charge electricity audit	Denmark
	Project 'Red-Hot' (element of stand-by campaign)	Denmark
	The 'A' campaign 1999	Denmark
	Promotion campaign for efficient ventilation	Denmark
	Information campaign (2001)	France
	Local energy information centres (Espaces Info Energie, EIE)	France
	Audits ("Aides a la decision")	France
	Energy audits in industry	Korea
	Energy audits in buildings	Korea
	Energy Efficiency Rating Labelling	Korea
	Information centres in local region	Sweden
	Information and education programme 1998-2002	Sweden
Economic	Criteria adopted for the evaluation of primary energy savings in end-uses	Italy
	EE Certificates	Italy
	Rebate programme for highly efficient electric inverters	Korea
	Financial incentives for DSM	Korea
	Energy premium scheme households	Netherlands
	Energy Investment Reduction (EIA and EINP)	Netherlands
Voluntary Agreements	Canadian Industry Program for Energy Conservation (CIPEC)	Canada
	Voluntary Agreements	Korea
	Voluntary Agreements on Industrial energy Conservation 1990 - 2000	Netherlands
	Eco-energy	Sweden
Combined Policy Measures	Rebate programme for household appliances	Belgium
	STEM programmes	Sweden

1. INTRODUCTION

Experts from Sweden, the Netherlands, Korea, Italy, France, Denmark, Canada and Belgium worked together to prepare this evaluation *Guidebook*, which contains two volumes: this volume (I), which deals with evaluation theory and advises on how to conduct evaluations for five types of policy measures and programmes, plus volume II which covers the country examples. This introductory chapter presents a general overview and taxonomy on policy measures. Section 1.3 introduces the seven key analytic elements of evaluations that are used in the following chapters on evaluating the selected types of policy measures.

1.1 Overview of the Energy Efficiency Policy Measures and Programme Evaluation Guidebook

1.1.1 Objectives

The International Energy Agency Demand-Side Management Programme (IEA DSM) has prepared this *Guidebook* to provide practical assistance to administrators, researchers, and policy makers who need to plan assessments and to evaluators who carry out evaluations of energy efficiency programmes. The *Guidebook* has been specifically designed to meet evaluation needs that result from recent developments in the practice of energy efficiency programmes, as well as in their policy context. These developments include the following examples.

Governmental Commitments to Emission Reductions. Many governments around the world, including those of most European countries, have signed the Kyoto Protocol to the United Nations Framework Convention on Climate Change (UNFCCC). Signatories to the Kyoto Protocol (that is now in force) are committed to achieving significant reductions in annual CO_2 emissions. Virtually all scenarios created by national and international energy and environmental agencies identify increased efficiency in the buildings, transportation, and industrial sectors of the developed economies as the key mechanisms for achieving those emission reductions.[1] Given their importance to international environmental policy, the results of energy efficiency programmes need to be estimated as rigorously as possible, using broadly accepted and transparent methods.

Proposed EU Directive on Energy End-use and Energy services[2]. In this proposal the European Commission sets a target of 1% for average annual energy savings by the Member States. This target is expressed as an amount of energy that should be saved, attributable to energy efficiency measures by final customers (end-users). In order to implement and meet this target countries will have to quantify and verify the effect of the measures taken (or to be

[1] See, for example, International Energy Agency. (2002). *Beyond Kyoto: Energy Dynamics and Climate Stabilization.* Paris: OECD, pages 40-58.
[2] Proposal for a directive, European Parliament and the Council on Energy End-use, Efficiency and Energy Services, Brussels, 10.12.2003 COM(2003) 739 final

taken), using the guidelines given in the proposal. These guidelines will allow quantifiable savings to be attributed to all the energy services, energy efficiency programmes and other measures that have been taken, including measures taken as early as 1991. Also included are measures such as energy taxes, building codes and information campaigns.

Proliferation of Energy Efficiency Programmes and Policy Measures. Early government and utility companies' energy efficiency programmes made considerable use of economic incentives such as rebates, investment subsidies, and free services to encourage end-use customers to purchase energy-efficient equipment and construction designs. Evaluation of these programmes focused on the characterisation of participant response, measurement of energy savings among participants, and attribution of their actions to the programme, among other potential influences. Over the past 15 years, the variety of policy measures employed by energy efficiency programme sponsors has vastly proliferated. For example, many countries or groups of countries operate equipment-labelling programmes. These programmes develop voluntary agreements with manufacturers to produce high-efficiency appliances and other equipment to negotiated standards, and then provide marketing support for those items through labelling and brand-building campaigns. In evaluating these programmes, the categories of participants and non-participants are not relevant, at least among end-users. The same applies to the development and enforcement of building codes and minimum equipment efficiency standards, which address the behaviour of the full population of market actors on the supply side of energy technology markets. Evaluation practitioners have developed methods to characterise the effects of such programmes, and programme sponsors should be aware of these efforts.

International Scale of Programme Efforts. Many energy efficiency programmes now operate on an international scale. The European Energy Label Programme, which applies to domestic appliances, operates in all EU Member States. Most Member States have also implemented the Green Lights Programme to promote the use of efficient lighting in commercial buildings. The ENERGY STAR labelling programme for office equipment, originally developed by the United States Environmental Protection Agency (US-EPA), has also been adopted by many countries in Europe. Finally, the Kyoto Protocol itself specifies three international programmes: Joint Implementation (JI), the Clean Development Mechanism (CDM), and Emissions Trading (ET). Differences among local markets, programme adaptations, and regulatory schemes will probably preclude the application of the identical evaluation methods to these programmes in each country where they appear. However, continued international cooperation and learning will be supported by the use of a consistent set of research questions and analytical frameworks applied to specific programme types.

Increasing Variation in Evaluation Objectives and Target Levels. The proliferation of energy efficiency programme approaches and organisational settings has been accompanied by increased variation in the evaluation objectives. Until fairly recently, evaluation was often oriented towards developing an estimate of energy savings attributable to a programme that was defensible in regulatory proceedings. Evaluations that meet this standard are expensive (but justified) where programme responsibility is narrowly assigned to one organisation, where budgets are relatively large, and programme cost recovery depends upon meeting quantitative energy savings goals. Today, many energy efficiency programmes operate well outside this organisational framework. Their sponsors may find an evaluation that does not cover all aspects of programme performance, or that results in savings estimates with a much larger confidence interval, to be acceptable for planning and management purposes. A useful framework for planning evaluations must therefore relate methods not only to the nature of the programme, but also to the sponsor's evaluation objectives, need for precision and accuracy, and budget. These last factors can be referred to as the sponsor's *target level* for the evaluation. This factor is included as one of our key elements for evaluation.

In light of these developments, the experts responsible for developing this *Guidebook* decided that it should meet the following objectives.

1. Provide guidance for the evaluation of a broad range of energy efficiency programmes currently offered by governments, energy companies, and other sponsors, with specific approaches for each major type of programme or policy measure.
2. Focus on providing guidance in matching research questions and methodological approaches on the one hand, to programme type and target level on the other.

The following sections discuss how these objectives are addressed in the *Guidebook.*

1.1.2 Intended Applications of the Guidebook

Systematic approach to methods selection. This *Guidebook* is designed to help programme managers, analysts, and planners carry out the following tasks in evaluating specific energy efficiency programmes.

- State evaluation research questions and hypotheses in relation to broader policy objectives and benefit-cost frameworks.
- Identify appropriate analytical methods.
- Specify data requirements and effective data collection methods.
- Set realistic expectations among the stakeholders of the evaluation regarding the nature and practical value of results to be delivered and the expected precision of quantitative estimates of programme effects.
- Set appropriate schedules.

In practice, the approach to these tasks will be shaped by many situational factors. Among the most important are:

- The nature and objectives of the programme.
- The current stage in the programme's lifecycle.
- The policy and/or regulatory framework for the application of evaluation results.
- Available budget or, related to that, the relative priority placed upon the evaluation's comprehensiveness and accuracy by the responsible authorities.

This *Guidebook* presents a systematic approach to selecting evaluation methods for a given situation, as defined by the factors listed above.

Guidance to appropriate sources. As the variety of policy measures, as well as energy efficiency programme types and objectives have proliferated, so too have the strategies used to evaluate them. Two strategies are pursued in order to address the need for more detailed guidance. The first is to produce this *Guidebook*, which contains seven key elements for the major types of measures or policies. The authors feel that each evaluation should focus on:

1. Policy measure theory that is used for developing and implementing a measure.
2. The choice and specification of indicators showing the success of a measure.
3. The baselines for the selected indicators.
4. Assessment of outputs and outcomes of the policy measure.
5. Assessment of energy savings and emissions reductions and other relevant impacts of the policy measure.
6. Calculation of cost, cost-efficiency and cost-effectiveness.
7. The choice of level (for evaluation efforts).

The choice of these key elements is argued in detail in Section 1.3. There are also extensive references to full evaluation reports and technical materials that can provide a high level of detailed guidance with regard to the highly specific methodological questions that so often arise in carrying out energy efficiency programme evaluations[3].

The second objective is to provide case summaries of recent evaluations in the participating countries and place the methods used within the context of the specific type of programme evaluated, its regulatory setting, the stage of programme development, and the scale of the evaluation effort[4].

1.1.3 Scope of Programmes and Policy Measures Covered[5]

This *Guidebook* covers the evaluation of programmes that:

- Target the reduction of energy use in the buildings and industrial sectors through increased end-use efficiency, reducing the level of energy services consumed, or some combination thereof.
- Promote customer adoption of commercially available energy-efficient equipment, as well as the adoption of efficient practices in design, specification, and construction.
- Promote the manufacture, distribution, and merchandising of energy-efficient equipment, as well as efficient construction practices among firms and individuals on the supply side of the relevant product and services market.

This encompasses a very broad range of efforts. It includes rebate programmes for energy-efficient equipment, rebate programmes for the adoption of efficient practices in new construction and renovation, tax-based incentives for end-use reduction or investment in efficient equipment, appliance labelling programmes, general customer information campaigns, vendor and design professional training, voluntary agreements, and energy audits.

---

[3] See references and reports at www.dsm.iea.org

[4] Case examples on evaluation are presented in Volume II

[5] Definitions of the terms ‘programme’ and ‘policy measure’ are discussed in detail in Section 1.2.

The experts use the following structure for policy measures and programmes in this *Guidebook*:

1. Regulations.
2. Related information.
3. Economic incentives.
4. Voluntary agreements.
5. Combinations of policy measures and programmes.

For more information please refer to Section 1.2.

The following kinds of programmes are <u>excluded</u> from consideration in the *Guidebook.*

- Programmes to reduce emissions generated by energy (and particularly electricity) supply activities, such as environmental regulation of thermal energy plants, emissions trading, or differential pricing of input fuels.
- Programmes that promote the development of renewable energy resources.
- Research and development programmes for new energy-related technologies.

The experts decided not to cover these kinds of programmes because their evaluations raise substantially different sets of issues from those encountered in evaluating programmes to promote market-ready energy efficiency technologies.

1.1.4 Types of Evaluations Covered

Focus on outcomes and energy impacts. The *Guidebook* focuses on the evaluation of policy measures and programme outcomes and energy impacts (for more detail see Section 1.3.2). Following the conventions of evaluation literature,[6] the term 'outcomes' includes the effects of a technological as well as a behavioural programme. These effects include changes in customers' or vendors' awareness levels of energy efficiency opportunities, changes in attitudes towards the adoption of energy-efficient operation and maintenance of equipment and facilities, changes in the availability and price of energy-efficient equipment, and changes in the market share of energy-efficient models of various types of equipment etc. All of these can be related, through variably complicated causal chains, to reduction in energy use. For our purposes, 'impact evaluation' refers to the estimation of energy savings and CO_2 reductions attributable to the programme.

Policy measures and programme versus project focus. This *Guidebook* is designed to support evaluations of programmes that target reduced energy consumption among groups of end-use customers or facilitators or intermediate organisations. It is not designed to deal with energy improvement in individual facilities. The *International Performance Measurement*

---

[6] Technopolis France 2001, Evaluation and Impact Analysis of RTDI Programmes, presentation by DG Research, 31 July 2001;
European Commission 1999, Paper I Defining criteria for evaluating the effectiveness of EU environmental measures;
Dutch Ministry of Finance, 2003, Guidance evaluation research ex post (in Dutch);
European Commission Agriculture Directorate-General, January 2002, Guidelines for the evaluation of LEADER+ programmes (document VI/43503/02-REV.1);
Gretchen B. Jordan, August 19, 2003, Developing Logic Models and Using Them to Define a Balanced Set of Metrics, handouts at IEPEC workshop, 2003

and Verification Protoco (IPMVP)l[7] provides extensive guidance on estimating energy savings associated with various kinds of energy efficiency improvements in single, relatively large facilities. The techniques described in the IPMVP can be viewed as one component of the tools and methods available for conducting programme evaluations.

Output and impact versus process evaluation. *The Guidebook does* ***not*** *cover process evaluations,* i.e.: evaluations designed only to assess the degree to which programmes are being operated efficiently in regard to their stated objectives and organisational structure. This is not to imply that impact evaluations, especially as they are defined here, cannot furnish a great deal of useful information on which to base recommendations and proposals for programme improvements.

1.1.5 Structure of the Guidebook

The remainder of this chapter lays out a basic framework of selecting methods for evaluating specific energy efficiency programmes, and includes:

- A taxonomy of energy efficiency policy measures that is useful for quickly identifying the relevant evaluation issues associated with a given programme.
- Identification of seven key analytic elements that need to be addressed in virtually all energy efficiency programme evaluations.

Chapters 2 through 6 apply this framework in providing guidance to the types of policy measures identified in the taxonomy, as well as to programmes that combine several different policy measures. Each chapter is structured according to the seven key analytic elements:

- Policy measure theory used.
- Specification of indicators for the success of a measure.
- The baselines for the selected indicators.
- Assessment of outputs and outcomes.
- Assessment of energy savings and emissions reductions and other relevant impacts.
- The calculation of cost, cost-efficiency and cost-effectiveness.
- The level of evaluation effort.

Chapter 7 lists the conclusions and includes experiences from the case studies and the country reports as presented in Volume II.

A selection of case examples have been described in Volume II to allow wider access to present existing evaluation practice and experience often not available in English. The choice of examples has been determined by the need to illustrate certain critical aspects of evaluation theory and the present evaluation practice. They have not been selected as

[7] International Performance Measurement & Verification Protocol Committee, 2002. *International Performance Measurement & Verification Protocol: Volume I. Concepts and Options for Determining Energy and Water Savings.* Oak Ridge, TN: US Department of Energy, Office of Scientific and Technical Information.

examples of particularly good or poor evaluations. Furthermore, the focus of the case descriptions is the evaluation methodology and not the evaluation results.

1.2 Taxonomy of Energy Efficiency Policy Measures

1.2.1 Some Basic Definitions

To facilitate development of this *Guidebook*, as well as subsequent reporting of programme evaluation results, it was necessary to develop clear definitions of terms such as ‘policy’, ‘measure’, ‘programme’, and ‘project’. Many international and national organisations have compiled and catalogued information on energy efficiency programmes operating within their spheres of responsibility. These organisations include the United Nations Framework Convention on Climate Change, the European Union, and various national energy agencies.[8] The general objective of this work is to aggregate the various ongoing programme efforts in order to provide a more comprehensive view of such activities. Each sponsoring organisation uses a slightly different focus when defining policies and programmes.

Based on a review of this literature and our own understanding of common practice, the authors decided to use the following structure and adopt the following definitions for use in this report. Policies and measures are distinguished on the one hand, and programmes and projects on the other. The term policies is reserved for general (public) policies; e.g. Public Health Policy, Public Housing Policy, Environmental Policy, and Energy Policy. To implement a policy a government can choose between (policy) measures or (policy) instruments; e.g. fiscal arrangements, subsidies, permits, regulation etc. A programme is a market intervention, using a package of selected policy measures during a specific time period. Each programme is defined by a unique combination of programme strategy, market segment, marketing approach and energy efficiency measures(s). A programme is implemented by projects using inputs, which results in outputs.

> **Policy measure:** A specific type of political action or market intervention designed to persuade energy consumers to reduce energy use and encourage market parties to promote energy-efficient goods and services.
>
> **Programme:** An organised set of projects targeted towards defined market parties over a specific time period to achieve increased end-use energy efficiency or reduced use of energy services. A package of selected policy measures is used. This selection is based on a programme theory.
>
> **Project:** An organised set of activities to create output(s).

1.2.2 Range of Policy Measures Addressed in the Guidebook

International organisations and agencies have developed categories of policy measures to support assessments of aggregate levels of activity, comparison of policy outcomes, and the characterisation of activities in individual countries (see footnote 7). These categories are

---

[8] International Energy Agency. 2002. *Dealing with Climate Change: Policies and Measures in IEA Member Countries.* Paris. Subsidiary Body for Scientific and Technological Advice. 2002. *‘Good Practices’ in Polices and Measures Among Parties Included in Annex I to the Convention.* New Delhi: UNFCCC.

based largely on the means used to influence customer and vendor behaviour. In addition, most apply to the encouragement of emission reductions in the supply side as well as to a reduction in demand. The classifications put forward by the IEA, European Union, and UNFCCC, while not entirely uniform, are quite similar.

The UNFCCC guidelines on reporting and review cover the following types of policy or measure:

- Voluntary/negotiated agreements.
- Economic.
- Fiscal.
- Regulatory.
- Information.
- Education.
- Research.
- Other.

Figure 1.1 displays the list of policy measure types and subcategories developed by the IEA in the 2002 edition of its *Dealing with Climate Change* publication. This *Guidebook* addresses evaluation of the subcategories shown in **bold**.

Figure 1.1 List of Policy Measure Types and Subcategories from the IEA's 'Dealing with Climate Change' (2002)

Type of Policy Measure	**Subcategories**
Fiscal	**Taxes: taxes, targeted tax exemption, tax reductions, tax credits** **Fees/charges, refund systems** **Incentives/subsidies: transfers, grants, preferential loans, preferential funds,** feed-in tariffs
Tradable Permits	Emissions Trading[9] Green Certificates Project-based Programmes (Including CDM and JI)
Regulatory Instruments	**Mandates/standards** Regulatory reform
Voluntary Agreements	**Strong: monitored emission-reduction quotas with penalties or threat of stronger regulation for non-compliance** Weak: no monitored emission-reduction quotas
Research, Development and Demonstration	Research programmes Technology development Demonstration projects
Policy Processes and Outreach	**Advice/aid in implementation** **Consultation** **Outreach/information dissemination** Strategic planning and institutional development

[9] Emissions trading could be used as an economic incentive for energy efficiency investments. Italy, for example, is developing a system of tradable 'white certificates' to support the implementation of a national energy efficiency effort to be implemented by gas and electricity distribution companies. However, no programme experience is currently available for this approach.

1.2.3 Definitions of Policy Measures Used in this Guidebook

The experts discussed several options for structuring a list of policies, policy measures and programme types, based on literature and common use in the participating countries. For example, energy audits could be considered a separate type of policy measure or included in the broader type of 'information', financial incentives could be separated from fiscal, etc.
Figure 1.2 shows the list of policy measure types and subcategories developed by the experts involved in producing this *Guidebook*. This list includes four basic types and a group of 'combinations of policy measures'. This is not to be seen as *the* final list of policy measures, but as a good list from the viewpoint of evaluation. Each list contains certain restrictions, e.g. Subcategory 2.4 'energy audits' includes voluntary audits as well as mandatory audits, although there are major differences between these two. The following sections present definitions for each type and its subcategory.

Figure 1.2 List of Policy Measure Types and Subcategories, Developed for this Guidebook

Type of Policy Measure		**Subcategories**	
1	Regulation	1.1	Building Codes and Enforcement
		1.2	Minimum Equipment Energy Performance Standards
2	Information	2.1	General Information
		2.2	Labelling
		2.3	Information Centres
		2.4	Energy Audits
		2.5	Education and Training
		2.6	Demonstration
		2.7	Governing by Example
3	Economic	3.1	Project or Product-related Subsidies (rebates)
		3.2	Targeted Taxes, Tax Exemption, Tax Credits
		3.3	Financing Guarantees
		3.4	Third-party Financing Facilitation
		3.5	Reduced-interest Loans
		3.6	Bulk Purchasing
		3.7	Grants
		3.8	Technology procurement
		3.9	Certificate trading systems
4	Voluntary Agreements	4.1	Industrial Companies
		4.2	Energy Production, Transformation and Distribution Companies
		4.3	Commercial or Institutional Organisations
5	Combinations		

Policy Measure Type 1: Regulations. In this *Guidebook*, the term 'Regulations' refers to laws and implementation regulations that require certain devices, practices, or systems design to improve energy efficiency. The most common forms of regulations are:

- ***Building Codes.*** In some countries and regional jurisdictions, commercial and residential building codes contain provisions specifying required physical or performance characteristics for buildings or building subsystems.

- ***Minimum Energy Performance Standards.*** Minimum energy performance standards (MEPS) apply to energy-using devices such as domestic appliances, household electronics equipment, office equipment, transformers, electric motors, and packaged heating, ventilating, and air conditioning (HVAC) equipment. These standards generally contain two parts: the first states a minimum performance standard in terms that are relevant for the particular device, and the second specifies the testing procedures used to estimate or classify the energy efficiency of the subject devices or materials.

Policy Measure Type 2: Information. This policy measure is designed to:

- Increase the various parties' awareness and understanding of energy-efficient products and services, as well as their economic and environmental benefits.
- Persuade actors to change their behaviour towards adopting energy-efficient products and practices.
- Provide actors with the technical information they need to identify and adopt energy-efficient practices.

The evaluation literature identifies the following more specific types of information-based energy efficiency policy measures.

- ***General Information.*** These policy measures consist of paid advertising and public-relations campaigns designed to make consumers aware of the need to save energy, the means at their disposal to achieve this, and the consequences of not doing so.
  - ***Labelling.*** Most of the EU countries plus Japan, Canada, and the United States have adopted statutes and rules that specify product performance standards, testing procedures, and labelling procedures for energy-using products.
  - ***Energy Audits.*** Energy audits consist of a structured inspection of a facility to estimate energy use and identify opportunities for increasing energy efficiency. In some cases, it is the customers themselves who carry out the inspection using protocols developed by the programme manager. On-site observations are analysed to allocate metered facility energy use for specific end-uses, estimate savings associated with applicable efficiency measures, estimate the costs of those measures, and prepare investment analyses of those measures. Energy audits are designed to help facility owners overcome a number of common barriers to implementation of energy efficiency measures. These include reducing information costs, mitigating information asymmetries (by providing economic analysis of potential measures from a party with no financial interest in their implementation), and reducing perceptions of risk.
- ***Information Centres.*** Information centres package and disseminate relatively technical information on energy-efficient products and practices. These centres are generally designed to support the work of equipment vendors, engineers, and plant managers working in a relatively narrow market, defined by technology (e.g. lighting) or a specific branch of industry (e.g. food processing).

- ***Education and Training.*** Education concentrates on providing focused information on energy efficiency opportunities and the application of efficient technologies in particular end uses. Training focuses more on practical experiences.

- ***Demonstration.*** Once a new or improved technology for energy conversion or energy saving has been developed, this technology needs to be introduced into the market. Demonstration refers to the phase during which this new product or technique is tested in practice. This serves to generate information on the usefulness, costs and energy savings during real use or to demonstrate this product or technique to potential users or decision makers.

- ***Governing by Example.*** Governments (e.g. Belgium and the Netherlands) sometimes choose their own governmental buildings, appliances purchasing etc., for a programme to demonstrate energy savings.

Policy Measure Type 3: Economic. Economic policy measures offer the stakeholders financial incentives to adopt specified energy-efficient technologies in equipment replacement, remodelling, and new construction projects. The wide variety of financial incentives currently in use includes:

- ***Project or product-related subsidies (rebates).*** Rebates are offered for the documented use of specific products or construction techniques. Rebates are generally gauged according to the efficiency level and quantity of equipment installed.
- ***Targeted taxes, tax exemptions, and tax credits.*** Several European countries offer tax credits or accelerated depreciation for purchasing specified energy-efficient equipment. In some countries, partial exemption from fuel taxation is offered to facilities that meet agreed requirements for voluntary energy use reduction.
- ***Financing guarantees.*** Programme sponsors may offer credit guarantees to reduce risk premiums charged on loans to finance energy efficiency projects.
- ***Third-party financing facilitation.*** Third-party financing approaches, such as energy performance contracting, are used to finance energy efficiency projects. They often include a subsidy or credit guarantee that reduces the cost of the project to the customer.
- ***Reduced-interest loans.*** Some organisations offer reduced-interest loans to finance projects that incorporate specified energy-efficient technologies.
- ***Bulk purchasing.*** Organisations may aggregate large orders of energy-efficient equipment to receive favourable pricing from manufacturers. These price reductions are then passed on to the final customers purchasing the equipment.
- ***Grants.*** Amount of money given to an individual or organisation for a particular purpose.
- ***Technology procurement.*** A process through which a commodity, service or system is procured, and for which development of new technical solutions is essential in order to meet a specified requirement set by a buyer (or group of buyers). The development work may concern the product, system or the production process for which it is developed.

- ***Certificate trading systems.*** A system of green (or white) energy certificates is used to facilitate the market for renewable energy, energy savings or for energy efficiency improvements.

Policy Measure Type 4: Voluntary Agreements. Voluntary Agreements, as defined in this *Guidebook*, refer to policy instruments under which representatives of national or provincial governments enter into negotiation with facility owners or branch organisations to obtain a commitment to reduce energy consumption by a specified amount over a given time period. Such agreements frequently contain energy consumption monitoring protocols and provisions for technical assistance to participating facilities. The signatories generally face financial penalties for failure to meet their commitments under the agreement. This approach is often used in conjunction with targeted tax exemptions.

- ***Industrial companies.*** Voluntary agreements, negotiated agreements or long-term agreements between representatives of a government and a group of industrial companies, or an industrial association.
- ***Energy production, transformation and distribution companies.*** Voluntary agreements between representatives of a government and energy production, transformation and distribution companies (or their trade association).
- ***Commercial or institutional organisations.*** Voluntary agreements between representatives of a government and commercial organisations (e.g. financial organisations), institutional organisations (e.g. hospitals, schools) or even ministries (or their representative association).

Policy Measure Type 5: Combinations. Many contemporary energy efficiency programmes combine elements of two or more of the basic policy measures. There is a current trend towards combinations and packages of an increasing number of policy measures indicated in the subcategories. For example, efforts to promote energy-efficient appliances have featured label specifications (regulation), broad-based branding and merchandising efforts (information), consumer rebates for qualifying products (economic incentives), bulk purchasing by government entities (economic incentives), and support of design competitions to expand the supply of qualifying products.

1.2.4 Evaluation Case Examples

The experts from the selected participating countries conducted evaluations on policy measures and/or programmes. Figure 1.3 contains an overview of these examples, which are included in Volume II. The examples presented in the four basic types often include elements from other types of policy measures but they were not tangible enough to determine these as real combinations. It should also be kept in mind that these cases are examples of conducted evaluations and that the cases as such do not give a representative overview of all evaluations in a country. The following sections describe these examples to illustrate the key analytic elements of evaluations.

Figure 1.3 Overview of evaluation case examples by type of policy measure

	Policy type	Case examples	Country
1	Regulation	Building codes	Belgium
		Energy Efficiency Regulations for Residential Equipment	Canada
		Energy management scheme for large buildings	Denmark
		Minimum energy performance standards	Korea
		Energy Performance Standard (EPS) for houses	Netherlands
2	Information	Local energy efficiency information centres	Belgium
		Energuide for houses	Canada
		Energy labelling of small buildings	Denmark
		Free-of-charge electricity audit	Denmark
		Project 'Red-Hot' (element of stand-by campaign)	Denmark
		The 'A' campaign 1999	Denmark
		Promotion campaign for efficient ventilation	Denmark
		Information campaign (2001)	France
		Local energy information centres (Espaces Info Energie, EIE)	France
		Audits ("Aides a la decision")	France
		Energy audits in industry	Korea
		Energy audits in buildings	Korea
		Energy Efficiency Rating Labelling	Korea
		Information centres in local region	Sweden
		Information and education programme 1998-2002	Sweden
3	Economic	Criteria adopted for the evaluation of primary energy savings in end-uses	Italy
		EE Certificates	Italy
		Rebate programme for highly efficient electric inverters	Korea
		Financial incentives for DSM	Korea
		Energy premium scheme households	Netherlands
		Energy Investment Reduction (EIA and EINP)	Netherlands
4	Voluntary Agreements	Canadian Industry Program for Energy Conservation (CIPEC)	Canada
		Voluntary Agreements	Korea
		Voluntary Agreements on Industrial Energy Conservation 1990 – 2000	Netherlands
		Eco-energy	Sweden
5	Combined policy Measures	Rebate programme for household appliances	Belgium
		STEM programmes	Sweden

1.3 Key Analytic Elements of Policy Measure and Energy Efficiency Programme Evaluations

The following paragraphs describe the seven key evaluation analysis components that (according to the experts) apply to virtually every type of policy measure listed in Figure 1.2, to Energy Efficiency and to Demand-Side Management programmes.

The experts hold the opinion that these key analytic elements should be addressed in an evaluation. That does not mean that an evaluator has to carry out the work related to these elements from scratch. If a programme is well developed, the information on the first three elements – statement of the theory, specification of indicators and baseline – should already be available to the evaluators. In theory it does not matter whether one evaluates a policy measure, such as a building code, or a programme that is targeted to increase the good use of a building code through informational campaigns, subsidies and tools. But of course the list is not a recommended sequence for undertaking the evaluation. The best order always depends on the specific circumstances within policy instruments (or combination of instruments) and the emphasis within the evaluation. For example, selecting the correct level of evaluation effort (see key element seven) is relevant for the depth of the evaluation and thus also for the selection of indicators and the assessments.

These seven key analytic elements are:

1. Statement of policy measure theory.
2. Specification of indicators for evaluation.
3. Development of baselines for indicators.
4. Assessment of output and outcome.
5. Assessment of energy savings, emissions reductions and other relevant impacts.
6. Calculation of costs, cost-efficiency and cost-effectiveness.
7. Choice of level with regard to the evaluation effort.

Readers familiar with other energy efficiency programme evaluation guides[10] will note that the concepts featured prominently in those volumes – tracking system review, estimation of gross energy savings, estimation of net energy savings, free-rider effect and spillover, market transformation etc. – are included in these analytic elements ahead.

For the sake of clarity, the following section is restricted in the section headers and tables for ‘policy measures’ as programmes generally include more than a single policy measure. The examples show that, in most cases, programmes contain one of the policy measures as the key instrument and that the key analytic elements can also be used to evaluate EE and DSM programmes.

1.3.1 Key element 1: Statement of Policy Measure Theory

Overview. The policy measure theory statement provides the basic framework for the evaluation. During the development of the policy measure, a policy measure theory is selected (explicit or implicit) and information on this key element should be available to the evaluator. For the evaluation itself, the theory statement should include the following components:

A. Specification of the policy measure domain; and

---

[10] See, for example: SRC International AIS. 2001. *A European Ex-Post Evaluation Guidebook for DSM and EE Service Programmes.* Brussels: The European Commission;

Violette, Daniel. 1995. *Evaluation, Verification, and Performance Measurement of Energy Efficiency Programmes.* Paris: International Energy Agency;

Vine, E. and Sathaye, J. 2000. *Guidelines for Monitoring, Evaluation, Reporting, and Verification of Climate Change Mitigation Projects.* Washington D.C.: US Environmental Protection Agency.

B. Statement of policy measure effects hypotheses.

Policy measures that emphasise information transfer and voluntary agreements lengthen the causal chain between projects and the realisation of energy savings, extend the time period during which programme effects need to be monitored, and broaden the geographic area over which programme effects may unfold. Therefore, it is particularly important to state the policy measure theory clearly in evaluations of these information-oriented programmes. Moreover, a clear statement of the theory is required to identify appropriate methods and information sources for constructing baselines.

The 7 key analytic elements:

1. **Statement of policy measure theory**
2. Specification of indicators for evaluation
3. Development of baselines for indicators
4. Assessment of output and outcome
5. Assessment of energy savings, emissions reductions and other relevant impacts
6. Calculation of costs, cost-efficiency and cost-effectiveness.
7. Choice of level (evaluation efforts)

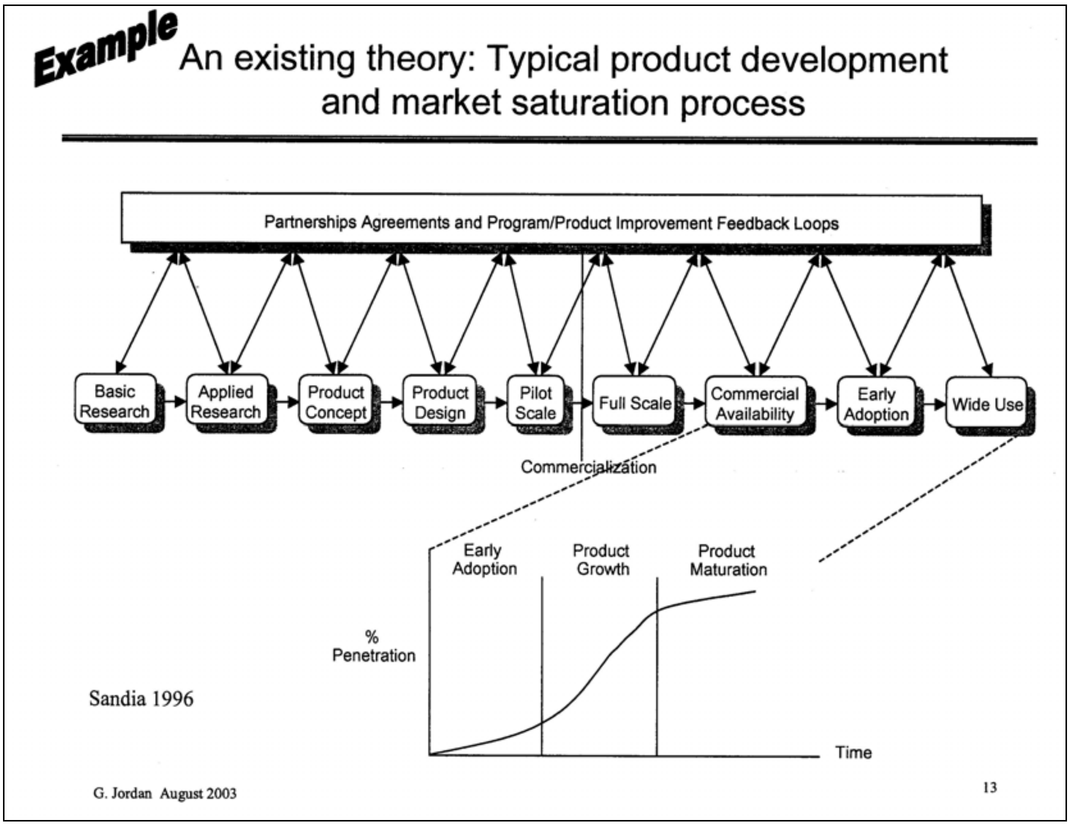
A. Specification of Policy Measure Domain. This component identifies the groups of market actors – end-users and inhabitants of the supply chain – who are likely to be affected by the programme. In order to structure evaluation efforts, the domain should specify the following.

- *End-user market segment.* This specification should go beyond the residential/commercial/industrial breakdown to identify the specific customers targeted. Useful categories may include size of facility or organisation, level of energy use, commercial building type or industrial branch, recent purchasers of the kinds of equipment promoted by the programme etc.
- *Supply-side market segment.* The specification should identify the level of the supply chain targeted (or emphasised) by the programme, for example: manufacturers, distributors, installation contractors, architects and engineers, but also financial institutes, lawyers, consultants etc.
- *Participation status.* For programmes that have a defined participation status, participants will certainly be included in the domain. Moreover, for many programmes it is plausible to hypothesise that non-participants among the targeted end-user or supply-side segments will be affected. This corresponds to the notion of non-participant spillover.
- *Location.* From a geographic standpoint, the domain will certainly include individuals and organisations that maintain facilities or do business in the area in which the programme is offered. The question here involves whether it is appropriate to extend this area in evaluating programme effects. For example, construction firms that develop energy-efficient practices in response to a programme offered in one state or region may apply those skills in other regions where they are active. Similarly, managers of national corporations who attend training seminars in one area may apply what they learn in facilities located elsewhere.

Determining where to draw the boundaries on the policy measure domain will require judgement from the evaluator and programme sponsor. It may be possible to hypothesise a wide range of programme effects on individuals who are not directly exposed to programme activities. However, it will only be worthwhile to assess these effects if they are likely to be large enough to affect programme cost-effectiveness *and* if the evaluator is likely to be able to demonstrate a credible link between those effects and the programme activities. For many kinds of programmes, the findings of past evaluations, as well as the observations of programme operators, can provide useful guidance in establishing the programme domain for evaluation purposes.

Figure 1.4 shows an example of an existing theory for the typical product development and market saturation process. For the last three phases, from commercial availability of a product (the early adoption phase) to the wide-use phase, the S-curve is often used for the assumed increase of market penetration. These kinds of theories should be combined with a specification of market segments and geographic area.

Figure 1.4 Example of an existing theory

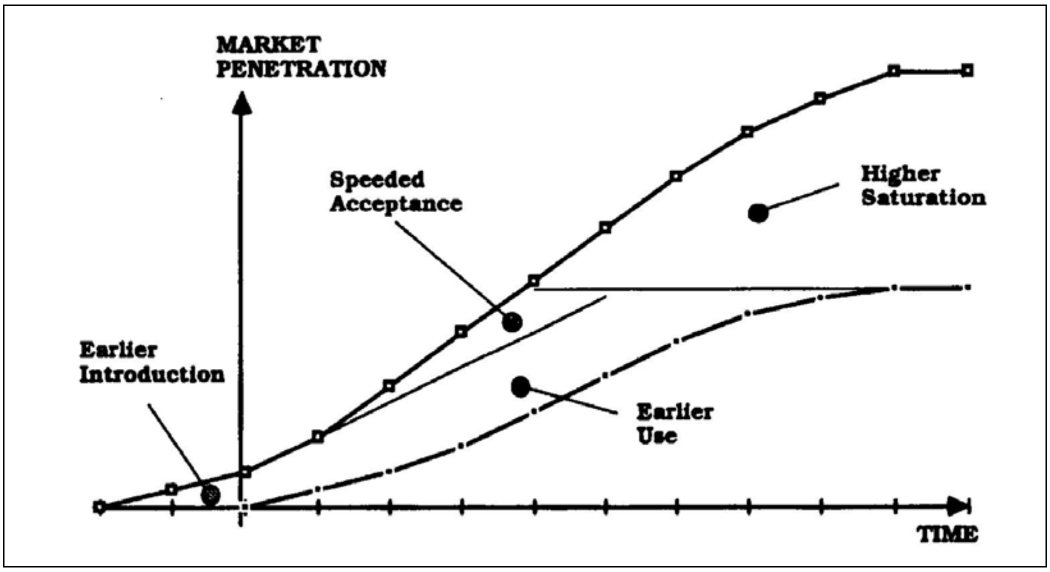


B. Statement of Policy Measure Effects Hypotheses. It is necessary to develop a clear statement of hypothesised effects in order to construct valid qualitative and quantitative measures of those effects. These hypotheses should also focus on how to allocate the impact that the various parties will have. A logical framework is often used to structure the hypotheses.

In the case of policy measures involving economic incentives, hypotheses concerning the effects on participants are straightforward: The availability of incentives or technical assistance helps customers overcome one or more barriers to implementing cost-effective energy efficiency measures. Those barriers might include the perception of performance risk, lack of available capital, lack of knowledge of the benefits and costs of the measures, high search costs, etc.

Hypotheses regarding programme effects on non-participants or for entire groups of market parties (where programme participation is not defined) are less straightforward. However, if the evaluator is interested in estimating energy savings realised by these groups, it is essential to specify the hypothesised causal path between the programme and those savings in as much detail as possible. For example, labelling programmes can hypothetically accelerate end-user adoption of efficient programmes in a number of ways: by raising awareness, reducing search costs, and providing manufacturers and retailers with co-marketing resources, to name just a few. Figure 1.5 gives a more detailed example of the theory of market transformation and sales as a direct indicator of customer demand for energy-efficient technologies.

Figure 1.5 Market transformation in terms of penetration

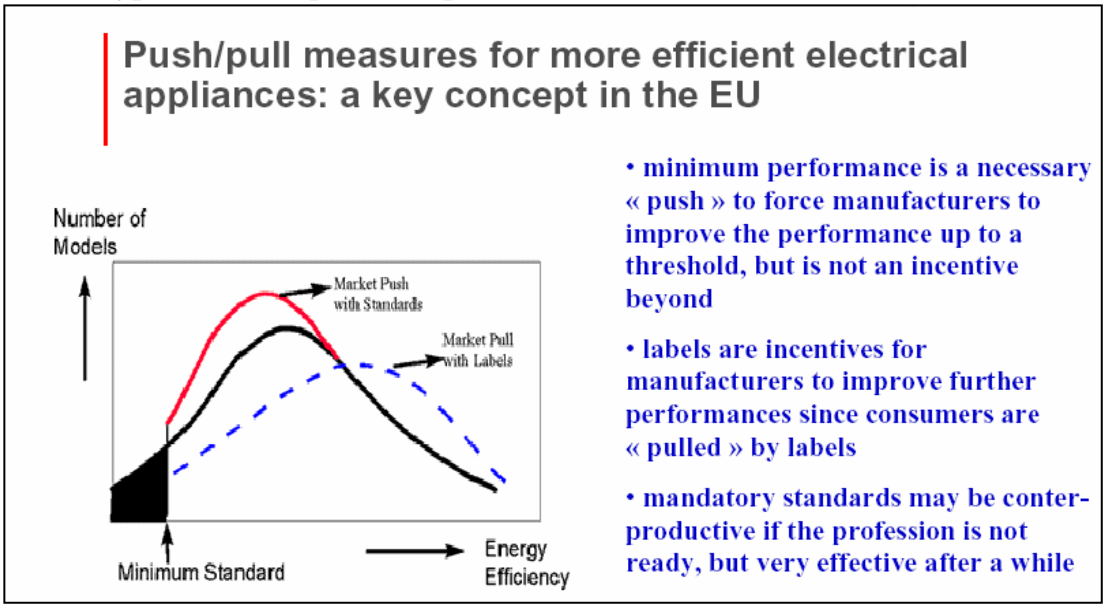


Source: Motiva 1999 page 126, based on Nilsson 1996

However, given the time it takes for consumers to assimilate brand information, as well as the prominence of other product considerations in model choice, it is reasonable to assume that a fair amount of time will elapse between the initiation of the education campaign and significant increases in the market share of efficient models. In the meantime, evaluators may need to specify and monitor indicators of intermediate results, such as increases in brand awareness and knowledge of efficient product attributes in order to demonstrate programme results. The point here is that hypotheses of programme effects may need to include the timeframe over which they are expected to become evident.

A final key function of programme effects hypotheses is to identify other important influences on end-user and supplier behaviour that need to be characterised and controlled when developing baseline values for programme effects indicators. To extend the labelling example, programme effect hypotheses would need to identify influences on end-user adoption of efficient equipment that are not addressed by the programme. These include, for example, the manufacturer's decisions regarding bundling of other desirable features in efficient models, manufacturer's decisions regarding incremental prices, the way in which the promotion of efficient models fits in with retailer strategies, and energy prices. These are also key considerations in establishing baseline equipment efficiencies and installation practices. Figure 1.6 gives an example of hypotheses on 'push and pull factors 'for efficient appliances

Figure 1.6 Hypotheses on 'push and pull factors'



Source: Ademe/Enerdata, 2003

Examples. Figure 1.7 illustrates components of specifying the programme theory statement for the policy measures addressed in this *Guidebook.* The columns contain examples for the domain specifications and for effects hypothesis. Examples of indicators on these effects hypothesis are presented later.

Figure 1.7 Examples of Elements of Programme Theory by Type of Policy Measure

Type of Policy Measure (Example)	**Domain Specification (Target Group Examples)**	**Effects Hypothesis (Examples)**
Regulation Building Code Enforcement	Builders	Builders increase frequency of using code-prescribed building methods. Building codes for new building also influence the existing buildings through retrofit
Information Programmes Labelling Programme	Purchasers of products subject to labelling Manufacturers and retailers of these products	Awareness of energy-efficient equipment increases among targeted consumers. Manufacturers increase share of efficient models in catalogues.
Residential Energy Audits Programme	Homeowners in programme area	Participants increase level of awareness and knowledge of efficiency opportunities in their homes. Participants implement targeted improvements more frequently than non-participants.
Economic Incentives Equipment Rebate Programme	Purchasers of equipment covered by programme Vendors and installers	Market share of efficient models increases. Proportion of vendors and installers promoting supported equipment increases.
Voluntary Agreements Industrial Programme	Owners of facilities in sectors covered by agreements	Owners and managers increase awareness of efficiency opportunities. Owners increase adoption rate for efficient production technologies.

1.3.2 Key element 2: Specification of Indicators

> *The 7 key analytic elements:*
> 1. Statement of policy measure theory
> 2. **Specification of indicators for evaluation**
> 3. Development of baselines for indicators
> 4. Assessment of output and outcome
> 5. Assessment of energy savings, emissions reductions and other relevant impacts
> 6. Calculation of costs, cost-efficiency and cost-effectiveness.
> 7. Choice of level (evaluation efforts)

The programme theory statement provides the basis for specifying indicators. An indicator is the evidence of information that represents the phenomena researched. Indicators help to gain knowledge of the results. They are measurable or observable: they can be seen (e.g. observed behaviour), heard (e.g. participant response), or read (e.g. agency records).

Based on existing literature (e.g. EC 1999, Dyhr-Mikkelsen, 2003, Dutch Ministry of Finance, 2003) and common practice, the experts agreed on the following structure of indicators relating to:

- Input (e.g. man-hours).
- Output (e.g. agreements with producers).
- Outcome (e.g. producers that comply with the agreement).
- Impacts, specified for energy savings and emission reductions.

In addition to these (examples of) physical indicators there are also an equal number of cost indicators. These are costs, cost-efficiency (euro/outcome), and cost-effectiveness (euro/impact, for example CO_2 shadow price). These cost indicators are used to compare the relative performance of different policy measures or packages of policy measures (for more details see Section 1.3.6).

Inputs refer to human and financial resources as well as other resources required to support a policy measure or programme. Partnerships and alliances might also be included here.

Outputs are the products, goods and services produced through programme activities and projects. Outputs are under the control of the management.

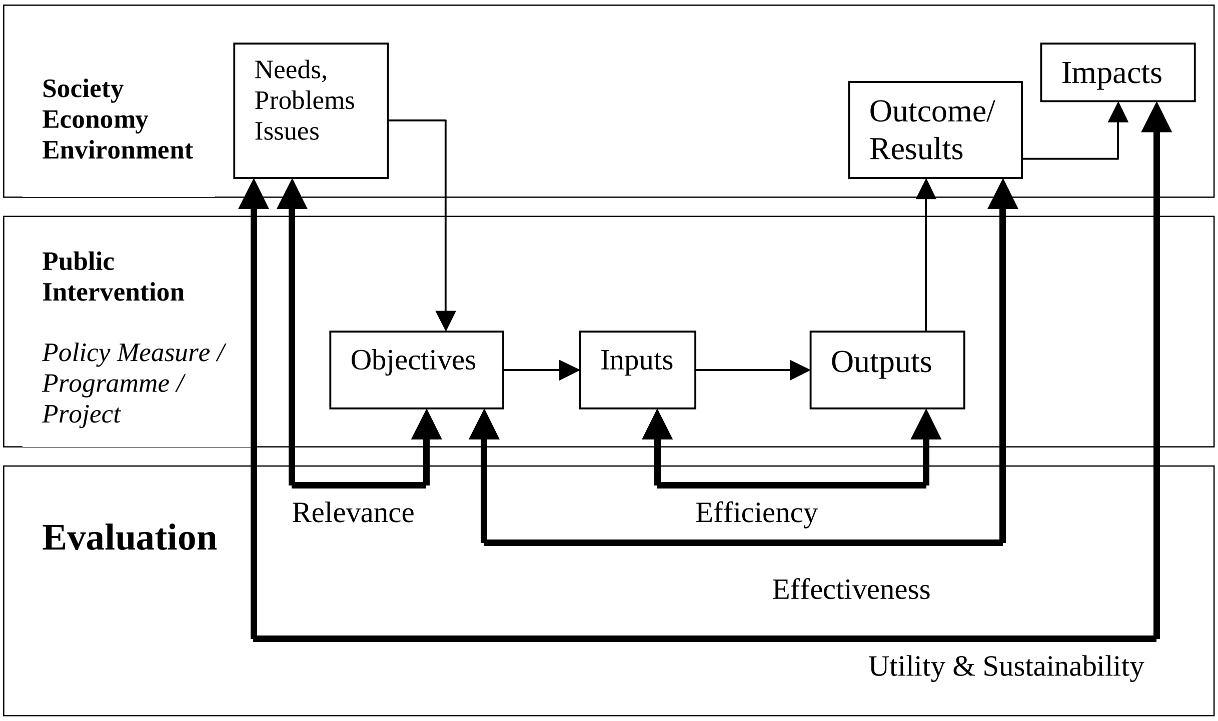
Outcomes or results refer to changes that occur from the use of outputs. Outcomes may be intended and unintended: positive and negative. Outcomes fall along a continuum from immediate (initial, short-term) to intermediate (medium-term) to final outcomes (long-term), and are often synonymous with impact.

Impacts are the social, economic, civic and/or environmental consequences of the policy. Impacts tend to be longer-term elements and so may be equated with goals. Impacts may be positive, negative and/or neutral: intended or unintended.

Figure 1.8 presents an overview of the relationship between the indicators and the use of terms such as efficiency and effectiveness. The experts see this as the preferred evaluation framework and so the approach in this *Guidebook* is consistent with this framework.

Figure 1.5 Preferred Evaluation Framework

Evaluation framework for a normative evaluation



Evaluation Questions

- *Relevance*: To what extent are the objectives justified in relation to needs?
- *Effectiveness*: To what extents have the expected objectives been achieved?
- *Efficiency:* Have the objectives been achieved at lowest cost?
- *Utility & Sustainability*: Do the expected or unexpected effects contribute to a net increase in social welfare and sustainability?

Adapted from European Commission 1999 and Technolopolis France 2001

The actual specification of the indicators must take into account not only their function in the evaluation but also prior knowledge about the availability and quality of data from various potential secondary sources and primary data collection methods. Often the programme management select indicators and collect information on these indicators during the implementation. The evaluator has to judge whether this selection and information is sufficient or whether (additional) indicators need to be selected.

The specification for each indicator should contain the following elements.

- ***Qualitative definition.*** The qualitative definition of each indicator conveys its general use in the analysis and its relationships to market conditions and programme effects identified in the programme theory. Indicators such as customer and supply-side actor awareness, market share, product pricing, vendor stocking (availability), and end-user purchase decision-making are important for programmes that promote efficiency technologies and measures that reach the customer in the form of free-standing products or equipment components. For programmes that promote efficient practices, indicators include the extent to which vendors have adopted practices such as the use of equipment sizing techniques approved by professional associations.
- ***Operational definition.*** Operational definitions provide the quantitative specifics of how each indicator will be developed. For example, the operational definitions of market share indicators will specify whether the indicator refers to the stock of equipment in place or the annual flow of sales, the technical definitions of ‘efficient’, the source of the market size measure to be used as a denominator, and the nature and source(s) of the measure of efficient technology adoption.

In reviewing potential sources of the data to be used in developing indicators of programme effects, the following characteristics need to be taken into account.

- ***Face validity.*** Do the data actually capture information on the programme or market attribute they purport to describe? For example, it may be more appropriate to use the number of vendors who have completed projects through a programme as an indicator of participation rather than the number who registered to participate.
- ***Accuracy.*** Do the data accurately represent the underlying facts of the situation? For example, it has been demonstrated fairly decisively that customers cannot accurately report over the phone the number of CFLs (compact fluorescent lamps) installed in their homes or whether they have purchased an appliance with a specific efficiency label. For these kinds of indicators validation is required by matching model numbers or through direct onsite observation by a trained data collector.
- ***Ability to be verified.*** Can key data, such as the amount of equipment installed in a given set of buildings, be verified by an independent body (if needed)?
- ***Absence of bias.*** The key consideration in assessing the potential bias of the measurement method is whether the full population of relevant supply-side establishments or customers’ reports to the relevant data collection agency is available to be sampled.

- ***Availability of historic data.*** Assessments of programme effects are greatly enhanced by the availability of historic data series going back to (or preferably prior to) programme inception.
- ***Replicability – availability of data on future sales or purchases.*** Given that the design of many programmes requires a multi-year effort for appropriate evaluation, any method selected will need to be replicable. For public data series, this means that the sponsoring organisations can be expected to continue their data collection efforts and provide public access to the results. For primary surveys, replicability requires the use of consistently developed sample frames and uniform, well-documented data collection methods over time.
- ***Comparability to other areas.*** Cross-sectional analysis is a powerful tool for assessing programme effects. The availability of data collected through consistent methods from a variety of states or regions is therefore a key advantage.
- ***Cost.*** Of course, the costs of primary data collection need to be taken into account in selecting data collection methods or in recommending the extension of current efforts.

1.3.3 Key element 3: Development of Baselines

Baseline estimation is the shorthand phrase for a group of methods used to address the research question:

'What would market actors who participated in (or who were exposed to) the programme have done in the absence of the programme?'

The 7 key analytic elements:

1. Statement of policy measure theory
2. Specification of indicators for evaluation
3. **Development of baselines for indicators**
4. Assessment of output and outcome
5. Assessment of energy savings and emissions reductions and other relevant impacts
6. Calculation of costs, cost-efficiency and cost-effectiveness.
7. Choice of level (evaluation efforts)

If a baseline has already been developed during the programme development and/or implementation phase management, the evaluator has to judge this and (if necessary) collect additional information. If not, he/she should develop baselines for the evaluation.

Economics and the social sciences yield a relatively small number of general strategies to address this question. The following methods are among those most frequently used.

Cross-sectional or quasi-experimental methods. These methods use observations of the equipment purchases, design decisions, or energy use of a group of market actors who were not exposed to the programme or who did not participate as the basis for characterising the baseline. Of course, for this approach to be valid, the comparison group must be similar in basic respects to the 'programme group'. Additionally, most cross-sectional methods employ statistical approaches to control the effect of differences between members of the programme and comparison groups along the dimensions that are likely to affect programme outcomes and impacts. These include differences in facility size, equipment inventory, local climate conditions levels of occupancy, levels of business activity etc.

Analysts who have used cross-sectional approaches have been concerned about controlling the effects of self-selection among programme participants. This concern arises from the observation, well established in economic literature, that the decisions to participate in a programme and to implement the supported improvements are often made simultaneously. Participants are distinguished from non-participants in that they perceive greater potential benefit from undertaking the improvements. In projecting the analysis results of a sample to the population, this systematic difference between participants and non-participants must be taken into account. Analysts have used a variety of multivariate linear and non-linear choice modelling methods to effect these adjustments.

The indicators of programme effect will generally include the adoption rate of efficient technologies, energy use, or changes in energy use over time. For many types of programmes, eligible customers who did not participate during the period under study may constitute an appropriate comparison group. This is the case, for example, in retrofit programmes oriented towards low-income households that occupy a fairly narrow range of a nation's or region's housing stock and who must generally register for programme services well in advance of receiving them. This enrolment process eliminates some typical concerns about self-selection of programme participants for voluntary programmes. For other types of programmes, comparison groups must be sought outside the programme area. This would be the case for 'market transformation' type programmes that broadcast their customer appeals. Finally, comparisons of programme effect can be made between regions differentiated by the intensity of programme effort or the particular type of programme design used to support a specific technology.

Historical or time-series analysis. In establishing programme baselines for evaluation, historical or time-series analysis involves developing a plausible historical narrative of programme effects, based on the relative timing of programme efforts and observed changes in programme effects indicators such as market share of energy-efficient equipment. The most convincing applications of the historical approach analyse changes over time in a broad range of programme-related indicators, such as measures of programme effort (amount of advertising or participants), product availability, average efficiency of available products, changes in provisions of codes and standards, price, market share, and market share within key segments. These trends can then be reviewed side-by-side to assess programme effects within the framework of the programme theory.

With a few rare exceptions, energy efficiency programmes (and the markets they target) do not allow for frequent enough observations of sales or market share to support formal time series analysis, as that term is used in econometrics and forecasting. Some established market share and technology forecasting methods have seen limited use in energy efficiency programme evaluation to forecast baseline market share. These include structured expert predictions (Delphi method), adoption process models, and learning theory. Examples will be discussed where appropriate.

Self-reporting of programme influence. In many cases, the decisions that energy efficiency programmes aim to effect, have many dimensions and involve complex processes within targeted organisations. For example, in estimating the baseline for a major commercial cooling system retrofit, the analyst will need to consider alternative basic technologies, sizing, efficiency of the equipment, and project timing. In assessing what the customer would probably have done in the absence of the programme, it will probably be necessary to interview the customer directly, and in some cases to interview multiple decision makers in large organisations. In analysing less complex purchases, such as efficient residential lighting and appliances, self-reports can be very useful in characterising programme effects on key customer decision elements, such as perceptions of product benefits and costs relative to conventional models.

Analysts have also made extensive use of self-reported methods for estimating levels of *free-ridership* associated with various programmes. Free-ridership refers to the portion of energy efficiency improvements made with programme support that programme participants would have implemented in the absence of the programme. Self-reported methods have also been used extensively to estimate spillover. Spillover refers to energy efficiency improvements that occurred due to programme influence, but were not directly supported through the programme with financial incentives or technical assistance. Among programme participants, spillover may occur when a facility owner has a favourable experience with an energy efficiency measure supported by the programme, and proceeds to implement the same kind of improvement in other facilities without programme support. Vendors who participate in a programme may decide to promote efficiency measures more broadly, based on favourable programme experience.

While methods based on self-reports are relatively straightforward to implement, they do have drawbacks. First, respondents are often tempted to exaggerate or underplay the influence that the programme had on their measure implementation decisions, depending on their perception of interviewer expectations and values. Second, in many cases, respondents often experience problems in recalling and reconstructing the circumstances around the purchase. Some of these problems can be mitigated to some extent by questionnaire design and through consultation with multiple sources for a given decision. However, these approaches cannot entirely eliminate the sources of potential bias in self-reported data.

Reference to codes and standards. For some kinds of policy measures, minimum product efficiency standards or building codes currently in force may serve as a reliable indicator of probable customer actions in the absence of the programme. This approach is often used in evaluating programmes to promote the inclusion of advanced efficiency features into new construction. The potential drawback to this approach is that significant time may have elapsed since the last revision of the relevant codes and standards. In that time, standard design and purchase practice may have changed to incorporate more efficient features than those included in the codes and standards. Also, a number of studies based on field observations of newly constructed facilities have found significant adoption of practices more efficient than those prescribed by the applicable codes, at least for some end uses.

Another increasingly used discussion concerns laboratory test results versus practice. For codes and standards the energy use is prognostic using laboratory tests under standard conditions. The practice could be quite different from this standard, and then one could question whether the energy savings should be calculated from the laboratory test.

Backcasting and other normative approaches. Some studies have used forecasts of technology adoption based on a combination of policy targets, expert opinion, surveys of customer preferences, and past experience as the basis for developing baselines. These 'backcasts' may be used in conjunction with baseline estimates developed through other methods to arrive at a suitable comparison for measures of programme effect such as changes in the market share of efficient equipment or changes in average energy utilisation indices for targeted facilities.

Only in rare cases are the results of any one of these approaches to assessing net programme effects on efficiency technology adoption found to be definitive on their own. Therefore it is best to plan to capture information to support at least two, if not all, of the approaches to baseline development in estimating net programme effects. A general, international, overview of baseline constructions for energy efficiency projects is presented in the report 'An initial view on methodologies for emission baselines: case study on Energy Efficiency' (IEA/OECD, 2000). Figure 1.9 shows typical approaches to developing baseline estimates for the kinds of policy measures covered in this *Guidebook.*

Figure1.9 Examples of Baseline Development Strategies by Type of Policy Measure

Type of Policy Measure/Example	**Example Programme Baseline Development Strategies**
Regulation Building Code Enforcement	Building code provisions covering the targeted building components and end uses.
Energy Audits Residential Programme	Non-participants' adoption of measures supported by the audits, properly adjusted for differences between the participant and non-participant groups.
Labelling Programme	Market share of qualifying equipment in areas not covered by the labelling programme. Historical trends in percentage of qualifying models sold by manufacturers and/or displayed by retailers.
Economic Incentives Equipment Rebate Programme	Market share of qualifying equipment in areas not exposed to rebate programmes, Non-participants' level of adoption of targeted technologies or end-use consumption, with appropriate statistical controls.
Voluntary Agreements Industrial Programme	Historical trends showing the improvement in energy consumption per employee or per unit produced, prior to the agreement. Owner reports of probable actions in the absence of the programme.

1.3.4 Key element 4: Assessment of Output and Outcome

Output. Programme activities and conducted projects produce products, goods and services that reach participating organisations and individuals. Indicators for outputs can therefore be specified for two groups:

- related to activities (what we did)
- related to participation (who we reached)

The 7 key analytic elements:

1. Statement of policy measure theory
2. Specification of indicators for evaluation
3. Development of baselines for indicators
4. **Assessment of output and outcome**
5. Assessment of energy savings and emissions reductions and other relevant impacts
6. Calculation of costs, cost-efficiency and cost-effectiveness.
7. Choice of level (evaluation efforts)

Output indicators related to activities are mostly quantitative: absolute (numbers) or relative (%) and refer to facilities, products, product development, workshops, media, onsite research, audits etc. Output indicators relating to participation are also mostly quantitative and refer to participants, organisations, clients, customers, users, etc.

Output indicators are often included in the operational objectives.

Figure 1.10 shows examples of output indicators relevant to the principal types of policy measures addressed in this *Guidebook.*

Figure 1.10 Examples of Output Indicators by Types of Policy Measure

Type of Policy Measure/Example	Example Output Indicators
Regulation Building Code Enforcement	Number of residences inspected and certified.
Information Energy Audits Residential Programme	Number of audits. Number of courses for energy auditors.
Labelling Programme	Percentage of equipment that contains a label. Percentage of qualifying models displayed with appropriate labels.
Economic Incentives Equipment Rebate Programme	Percentage of eligible facilities that participate in the programme. Market share of qualifying products.
Voluntary Agreements Industrial Programme	Percentage of facilities in the sector that sign the agreement. Percentage of signatories that comply with the agreement.

Outcomes or results are changes that occur through the use of outputs. Outcomes are the changes or improvements for individuals, groups and organisations that occur during or after the programme. So outcome indicators should represent the difference that the programme makes.

Outcomes often fall along a continuum from shorter to longer-term results. The fastest are the short-term or direct outcomes, which are often related to learning and change in knowledge and attitudes. Sometimes these immediate outcomes seem similar to outputs, as the programme has the most influence on the more immediate outcomes.

The next outcomes are the intermediates, occurring in the medium term. These are mostly related to behavioural changes, to decisions, or to actions taken as a result of the programme outputs.

The last group concerns the long-term (ultimate) outcomes. These are often synonymous with impact. This *Guidebook* restricts impact to the energy and environmental elements of a programme.

An information programme is used to illustrate the difference between outputs and outcomes. The output can be a leaflet that should make the reader aware (short-term outcome) of a problem. As a next step, the reader/consumer decides to gain knowledge or skills to change his/her situation (intermediate outcomes) and finally changes the situation (e.g. by installing a new boiler). The end result (the impact) is reduced energy use.

Outcome indicators are often included in the specific objectives.

Figure 1.11 shows examples of outcome indicators relevant to the principal types of policy measures addressed in this *Guidebook.*

Figure 1.11 Examples of Outcome Indicators by Types of Policy Measure

Type of Policy Measure/Example	**Example Outcome Indicators**
Regulation Building Code Enforcement	Saturation of code-required building features such as levels of insulation or efficient heating systems in annual cohorts of facilities.
Information Energy Audits Residential Programme	Percentage of audit recipients that implement recommended improvements.
Labelling Programme	Percentage of consumers who recognise and understand the message on the label. Market share of qualifying products.
Economic Incentives Equipment Rebate Programme	Percentage of facilities that use the equipment in the right way.
Voluntary Agreements Industrial Programme	Number of actions taken within the facilities in the sector that signs the agreement.

1.3.5 Key element 5: Assessment of Energy Savings, Emissions Reductions and Other Relevant Impacts

The 7 key analytic elements:

1. Statement of policy measure theory
2. Specification of indicators for evaluation
3. Development of baselines for indicators
4. Assessment of output and outcome
5. **Assessment of energy savings and emissions reductions and other relevant impacts**
6. Calculation of costs, cost-efficiency and cost-effectiveness.
7. Choice of level (evaluation efforts)

Energy impacts. The estimation of programme energy impacts generally proceeds in two stages.

- ***Estimation of gross energy impacts.*** Gross energy impacts refer to the change in energy consumption and/or demand that results directly from programme-related actions taken by end-users that are exposed to the programme, regardless of the extent or nature of programme influence on customer behaviour. This is the physical change in energy use after taking into account factors beyond the customer or sponsor's control. These include weather conditions, broad economic conditions that affect production levels and energy prices, changes in occupancy etc. Estimates of gross energy impacts always involve a comparison of changes in energy use over time among customers who installed measures and some baseline level of usage. Baselines may be developed from energy use measurements in comparable facilities, codes and standards, direct observation of conditions in buildings not addressed by the programme, or facility conditions prior to programme participation.
- ***Estimation of net energy impacts.*** Net energy impacts refer to the percentage of the gross energy impacts that is attributable to the programme. For transactional programmes, estimating net energy impacts involves the application of *free-ridership* and *spillover* concepts. Free-ridership refers to the portion of energy savings that participants would have achieved in the absence of the programme through their own initiatives and expenditures. Spillover refers to the adoption of measures by non-participants and by participants who did not claim financial or technical assistance for additional installations of measures supported by the programme. For programmes in

which participation is not well defined, the concepts of free-ridership and spillover are less useful. Estimating net energy impacts for these kinds of programmes generally requires the application of cross-sectional and historical methods to sales or market share data in order to estimate net levels of measure adoption.

Method selection for estimating gross energy impacts. Analysts have developed a broad range of strategies and techniques to estimate the gross energy impacts of energy efficiency programmes. The selection of methods for a particular programme will depend upon many factors, e.g.: the type of policy measure (as per the taxonomy above), the stage of programme operation, end-use technologies targeted, baseline information available, targeted precision levels for the savings estimate, and budget. The following paragraphs briefly describe the principal energy impact estimation methods. Figure 1.12 shows the most common applications of these methods by programme type and targeted end-use technology.[11]

- ***Engineering methods.*** Engineering methods develop estimates of energy savings (in energy units per year or as a percentage of annual energy use at the facility or end-use level) based on technical information on equipment performance and assumptions concerning operating characteristics of the equipment or facilities in which it is installed. The most straightforward application of engineering methods involves using savings algorithms that summarise how energy use is expected to change due to installation of the energy efficiency measure. This approach is best applied to programmes that involve equipment retrofits or replacing failed equipment with efficient models. Engineering methods have also been applied to estimating the effects of more complex measures, such as the adoption of efficient new construction practices or HVAC equipment through the use of building simulation models. In these cases, the models are calibrated to baseline conditions concerning building size, equipment, construction, and occupancy. Average savings are then estimated by changing the model parameters that are affected by programme participation.
- ***Short-term monitoring to enhance engineering methods.*** In some cases it is desirable and cost-effective to supplement engineering methods with short-term or spot-monitoring of site conditions. The results are then used to calibrate key parameters in the energy savings algorithms. Commonly measured parameters include operating hours for lighting and HVAC equipment, wattage for lighting and HVAC equipment, and line temperature and pressure for various refrigeration and fluid applications.
- ***Bill analysis.*** Bill analysis applies a variety of statistical methods to measured facility energy consumption data collected both before and after the installation of measures to estimate gross energy impacts. These approaches are generally applicable to programmes that meet the following criteria.
  - Participation is well defined.
  - There are a relatively large number of participants (over 100).
  - Sufficient programme experience has been gained to form a group of participants with measures installed for at least one year.

[11] For more extensive discussions and descriptions of alternative methods for estimating gross and net energy savings, see SRC International, op. cit. and Vine & Sathaye, op. cit.

- Expected changes in energy consumption due to measures installed through the programme account for at least 5% of facility energy consumption, and preferably 10% or more.

In their simplest form, statistical models apply one of a number of regression analysis techniques to measured energy use data to control for variations in weather during the evaluation period compared to a standard meteorological year. The average change in weather-adjusted consumption from the pre- to the post-programme period is then estimated for all (or a sample of) participants, and for a sample of non-participants. The difference between the average change in energy use for the two groups represents the net programme effects.

This basic approach requires analysts to make a number of assumptions that are not tenable for many programmes. The most important are that the comparison group closely resembles the participant group along all dimensions that affect the change in energy use and that the measures installed are uniform. These assumptions seldom apply in commercial retrofit and replacement programmes, where both the participating facilities and the extent of measures installed vary widely. Analysts have therefore developed more complex models to take account of these variations. One general approach, known as conditional demand analysis (CDA), develops a regression model of energy consumption using detailed data on facility configuration and operation, as well as programme participation, as independent variables. A further variation, known as statistically adjusted engineering (SAE) analysis, uses an engineering estimate of programme savings based on site-specific data as an additional input variable. Finally, analysts have applied a variety of regression techniques in an attempt to correct for the effects of self-selection bias among participants.

- ***End-use metering.*** Energy savings can be measured for specific end-uses by directly metering the energy consumption of equipment affected by the programme. This type of metering is conducted before and after measure installation, in order to characterise the performance of the equipment under a variety of load conditions. The data are often standardised for variations in both operations and weather. This method of data collection is expensive. The most appropriate applications for end-use metering involve measure-expected savings that are large in the absolute sense, but relatively small in relation to total energy use for the facility in which they are found. This situation applies to HVAC upgrades in large commercial facilities and in most efficiency upgrades to major manufacturing and agricultural production systems.

Many evaluations use a combination of techniques to develop estimates of gross energy impacts. One of the most common integrative techniques involves the projection of relatively expensive end-use metering or short-term monitoring results from a small subsample to a larger sample – and thence to the population – using ratio-estimation procedures.

Figure 1.12 shows typical applications of the principal methods for estimating energy savings.

Figure 1.12 Common Applications of Basic Methods to Estimate Energy Impacts

Method	Typical Policy Measures	Typical End-Uses
Engineering	Economic Incentives: tax-related measures and rebates Information programmes: labelling Energy Audits	Residential Lighting and Appliances Commercial Lighting Industrial Motors
Engineering with building simulation modelling	Regulation: building codes Economic Incentives: tax-related measures rebates Information programmes: labelling	Commercial New Construction Residential New Construction Commercial HVAC
Engineering with monitoring	Economic Incentives: rebates Energy Audits Voluntary Agreements Regulation: building codes and equipment standards	Residential HVAC Commercial Lighting Industrial Motors C&I HVAC and Refrigeration Plug-load equipment
Bill Analysis	Economic Incentives: rebates Voluntary Agreements	Residential Shell and Heating Commercial Lighting & HVAC Residential New Construction Commercial New Construction
End-Use Metering	Economic Incentives: rebates Voluntary Agreements	Residential New Construction Commercial HVAC & Chillers C&I Lighting Industrial Motor Systems (Compressed air, pumps)

Net-to-gross ratios. Net-to-gross ratios can be used to estimate free-ridership and spillover: a factor that represents the net programme impact divided by the gross programme impact. This factor is applied to gross programme savings in order to determine the programme's net impact. The California Public Utilities Commission uses a default value of 0.8 and for some programmes more specific values e.g. for residential audits 0.72, for commercial information programmes 0.83 and for industrial new construction incentives 0.62[12]

Estimating emission reductions. Once energy impacts have been estimated, estimation of emission reductions usually involves the straightforward application of emission factors.

There is a great deal of international and interregional difference in emission factors, due to differences in the fuel type, or for electricity, in the age and fuel mix of local electric systems. A number of organisations, including the Intergovernmental Panel on Climate Change (IPCC), the UNFCCC and the US Environmental Protection Agency, have developed detailed methods for converting site-specific energy savings to estimated emission reductions.[13] Figure 1.13 gives an example emissions factor table. The UNFCCC has developed a greenhouse gas

---

[12] California Public Utilities Commission, Energy Efficiency Policy Manual version 2, August 2003, page 19

[13] Intergovernmental Panel on Climate Change. 2001. *Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories.* Geneva: United Nations Environmental Programme. Also California Climate Action Registry. 2002. *Climate Action Registry Reporting Online Tool.* Los Angeles: California Climate Action Registry.

inventory database that is available from: http://ghg.unfccc.int/. For most countries the annual National Inventory Reports contain the country-specific emission factors. The 2003 country reports are available from: http://unfccc.int/program/mis/ghg/submis2003.html.

Figure 1.13 Example Table Showing CO2 Emission Factors of Selected Fuels

Fuel	**Energy value**	Unit	**CO_2[1]**	unit
Electricity				
NSW, ACT	3.6	MJ/kWh	0.968	Kg/kWh
Victoria	3.6	MJ/kWh	1.467	Kg/kWh
Queensland	3.6	MJ/kWh	1.04	Kg/kWh
SA	3.6	MJ/kWh	1.109	Kg/kWh
WA	3.6	MJ/kWh	1.032	Kg/kWh
Tasmania	3.6	MJ/kWh	0.002	Kg/kWh
NT	3.6	MJ/kWh	0.756	Kg/kWh
Australia average	3.6	MJ/kWh	1.051	Kg/kWh
Canada average	3.6	MJ/kWh	0.22(3)	Kg/kWh
Natural Gas				
NSW, ACT			63.2	Kg/GJ
Victoria			58.9	Kg/GJ
Queensland			56.7	Kg/GJ
SA			57.6	Kg/GJ
WA			62.3	Kg/GJ
NT			54.6	Kg/GJ
Australia average			59.4	Kg/GJ
Black coal			89.4	Kg/GJ
Brown coal	9.7(2)	GJ/tonne	87.7	Kg/GJ
Briquettes	22.1 (1)22.3(2)	GJ/tonne	103.0	Kg/GJ
Coke	27.0(1)28.5(2)	GJ/tonne	117.0	Kg/GJ
LPG	25.7(1)26.6(2)	MJ/l	64.7	Kg/GJ
Propane	25.53(3)	MJ/l	1.53(3)	Kg/l
Aviation gasoline	33.1(1)	MJ/l	73.3	Kg/GJ
Petrol	34.2(1)(2) 34.66(3)	MJ/l	71.3	Kg/GJ
Jet fuel	36.8(1)	MJ/l	73.1	Kg/GJ
Kerosene	36.6(1) 37.68(3)	MJ/l	74.9	Kg/GJ
Heating oil	37.3(1)37.6(2)	MJ/l	74.9	Kg/GJ
Automotive diesel	38.6(1)38.4(2) 38.68(3)	MJ/l	74.9	Kg/GJ
Light fuel oil No 2	38.68(3)	MJ/l	2.83(3)	Kg/l
Heavy fuel oil No 6	41.73(3)	MJ/l	3.09(3)	Kg/l
Industrial/marine diesel	39.6(1)38.6(2)	MJ/l	74.9	Kg/GJ
Fuel oil			78.8	Kg/GJ
high sulphur	42.9(2)	MJ/kg		
low sulphur	44.5(2)	MJ/kg		
Town gas			58.7	Kg/GJ

(1) Source Greenhouse Challenge Office; (2)source DPIE; (3) Caddet Newsletter March 01

Source: http://www.aie.org.au/melb/material/resource/fuels.htm

Other relevant impacts. Policy measures are often not only targeted towards energy savings and emission reductions, but are also used to improve employment in general (or in a specific geographic area), increase (international) industrial competitiveness, decrease the use of materials, reduce water pollution etc.

1.3.6 *Key element 6: Assessment of Costs, Cost-Efficiency and Cost-effectiveness*

The final step in most evaluations is to prepare an analysis of the cost, cost-efficiency and cost-effectiveness of the programme(s) in question.

The 7 key analytic elements:

1. Statement of policy measure theory
2. Specification of indicators for evaluation
3. Development of baselines for indicators
4. Assessment of output and outcome
5. Assessment of energy savings and emissions reductions and other relevant impacts
6. **Calculation of costs, cost-efficiency and cost-effectiveness.**
7. Choice of level (evaluation efforts)

Costs relating to a policy measure or a programme refer not only to the direct costs but also to indirect costs such as changes in tax revenues, transmission capacity costs or profit lost. In most evaluations the costs are restricted to the programme costs. The costs are related to the input indicators.

Cost-efficiency is the ratio between the input and the outputs. One should keep in mind that the input cost also includes the cost of generating the products, services etc. (the outputs of the programme). In most cases it only makes sense to use cost-efficiency for projects; e.g. are the agreed number of workshops organised at the lowest cost? For a programme or policy measure the cost-effectiveness (ratio input-outcome) is more appropriate.

Cost-effectiveness calculations can be made from various perspectives: e.g. the participants, energy distributor (the utility), programme administrator, total resource cost and society[14]. National authorities with responsibility for programme funding and review generally specify the form(s) of the cost-effectiveness analysis to be applied. Most of these approaches are designed to estimate the net social benefits of the programmes in question. They may differ along the following dimensions: timeframe over which programme effects are tracked and counted; valuation of energy and emission benefits; inclusion and valuation of non-energy benefits (such as increased productivity, reduced health risks, increased employment, etc.); treatment of various kinds of programme costs; selection of discount rates.[15]

The most common cost-effectiveness indicator is the Net Present Value (NPV) of programme impacts over the lifecycle of those impacts. In its simplest form this is:

NPV = Σ benefit /(1+interest rate)
While the summation is taken over the years that the programme has impact

Whatever the particular form of the cost-effectiveness equations, they generally require a fairly uniform set of input data. Figure 1.14 shows the data required by most cost-effectiveness tests.

[14] For details see: *California standard practice manual: economic analysis of demand-side programs and projects* October 2001

[15] See Sebold F. et al. 2001. *A Framework for Planning and Assessing Publicly Funded Energy Efficiency.* San Francisco: Pacific Gas & Electric Company. Sections 2 and 8 for a discussion of policy issues encompassed in the selection of cost-effectiveness tests and formulae.

Figure 1.14: Components of Cost Effectiveness Calculations

Data Category	Typical Items Required	Comments
Policy Measure Benefits: Energy	Net programme energy savings	Need to take into account unit savings, measure lifetimes, decay factors.
	Avoided costs per kWh, kW, or unit of gas saved	Choice of cost index depends on regulatory structure of the market. Forecast of prices required for estimates of 'out year' savings.
	Avoided distribution and transmission costs	Estimates of avoided costs must be made in reference to market structure – e.g. market price for reserves.
Policy Measure Benefits: Environmental	Volume of emissions reductions	Requires model (could be simple) of local electric supply system.
	Unit value of emissions reductions	Clearing price for emissions trading could be used if market in place. Otherwise, estimates derived from environmental damage studies can also be used.
Policy Measure Benefits: Non-Energy	Volume of water and other non-fuel resource savings Unit value of non-fuel resource savings Non-energy benefits: increased productivity, increased safety, and accelerated collections.	Need to be able to demonstrate the association of the benefit to the programme, estimate magnitude, and assess monetary value. Otherwise, must be considered anecdotally outside the c-e equations.
Policy Measure Costs: Administrative	Administrative personnel and overhead costs Outsourced programme administration Marketing and promotional costs Measurement and evaluation costs	In most social cost-effectiveness schemes, incentives paid directly to customers or vendors are identified as transfer payments and are not counted as costs.
Policy Measure Costs: Incremental project implementation	Incremental costs of measures implemented as a result of the programme.	Requires cost estimates for both the energy-efficient measure and its baseline alternative.
Policy Measure Costs: Other	Measure-specific items, such as costs to properly dispose of used fluorescent ballasts and lamps or downtime for installation.	Judgement required regarding: whether the probable magnitude of these costs is sufficiently high to justify measurement expenses.

Cost-effectiveness and CO_2 emission reductions. Another viewpoint when looking at the cost-effectiveness is to take CO_2 emissions reductions as a reference. The benefits of the policy measure or programme, the energy savings, are not seen in terms of money, but are viewed in terms of avoided emissions. This results in costs per ton CO_2 that can be compared with calculated costs for other options. One should be aware of the fact that a policy measure often has CO_2 reduction as a main target, but that more targets are also set for that policy measure. One could question whether it is correct that all costs are then accounted to CO_2.

The costs per ton of CO_2 vary per country and type of policy or technology as well the type of energy user. In relation to the EU CO_2 emissions trade, a price range from 5-58 euro/ton CO_2 was indicated in 2000, while the price in the British trading system was 4-19 euro. Several national sources indicated mid-2002 prices of 2.5-9 euro. The Dutch Ministry of Economic Affairs expected (mid 2004) that it would be possible to buy CO_2 rights in Joint Implementation projects for around 6 euro/ton (including programme costs). These prices are much lower than average costs relating to energy saving policies. The governmental costs relating to the Dutch Energy Premium Scheme for households were estimated (in the year

2000) from 120-340 euro/ton CO_2. On the other hand the greenhouse abatement being achieved by the Australian NAEEEP (National Appliance and Equipment Energy Efficiency Program) was calculated to cost minus Aus$ 30/tonne of avoided carbon dioxide.[16]

1.3.7 Key element 7: Level of Evaluation Effort

The 7 key analytic elements:

1. Statement of policy measure theory
2. Specification of indicators for evaluation
3. Development of baselines for indicators
4. Assessment of output and outcome
5. Assessment of energy savings and emissions reductions and other relevant impacts
6. Calculation of costs, cost-efficiency and cost-effectiveness.
7. **Choice of level (evaluation efforts)**

Relationship of evaluation efforts to programme lifecycle. Evaluation activities can provide a great deal of value to programme planners and managers at each stage of programme development: from planning, through to inception and finalising in maturity. Figure 1.15 shows evaluation objectives and activities typically associated with various stages of programme development and implementation. The authors therefore argue that attention should also be paid to the key elements for evaluation during the programme planning and programme implementation phases. This will improve the quality of the programme and reduce the evaluation work (and associated costs) during programme revision or after the programme has ceased. Also the planning of monitoring actions and including the related costs in future programme costs will also ease the evaluation. One should keep in mind that, although it takes time to prepare and conduct an interview etc., this can often only start once the programme has been running for some time.

Figure1.15 Evaluation Objectives and Activities, per Programme Development Stage

Programme Stage/Evaluation Objective	Examples of Evaluation Activities
Programme Planning *(-12 months to launch)*	
Prepare programme theory hypotheses.	• Review evaluations of similar programmes; in own countries and others nations. • Interview programme managers, key market actors.
Select indicators of output and outcome.	
Agree on estimates of baseline.	• Estimate size of market; characterise segments. • Compile market share information: sales, saturation, manufacturers catalogues. • Interview market actors regarding typical practices. • Compile information on efficiencies of current technologies; trends in efficiencies.
Estimate energy savings.	• Develop engineering algorithms and estimates of parameters for basic savings calculations.
Estimate emission reductions.	• Compile local information on converting units of energy saved to emission reductions.
Forecast of cost-effectiveness.	• Compile cost data for measures and programme operations

[16] Achievements NAEEEP 2003, Australian Greenhouse Office, March 2004, page 4

Programme Stage/Evaluation Objective	Examples of Evaluation Activities
Startup ***(Launch to +12 months)***	
Select indicators of programme outcomes and impacts.	• Estimate value of key indicators on small-scale test basis.
Agree on estimates of baseline.	• Collect information on baseline-related conditions. • Develop detailed documentation on programme activities (to support potential time-series analysis). • Specify facility and measure-related conditions to be captured in programme tracking system.
Estimate energy savings. Develop preliminary cost-effectiveness estimates.	• Initiate tracking data collection and data quality control; estimate savings and costs.
Full-Scale Implementation ***(+12 to +36 months)***	
Select indicators for programme outcome and impact.	• Collect information on key indicators through customer and supply-side market actor surveys.
Develop estimates of baseline.	• Capture baseline information for *post hoc* evaluation. Could include: data on relevant conditions among comparison groups, in different regions, over time in the same region. Alternatively, self-reports from participants and non-participants.
Estimate energy savings.	• Collect information to estimate gross change in energy use by participants or target group for information and regulatory programmes. • Combine baseline and gross savings estimates to develop estimate of net savings. • Apply factors to estimated net energy savings to estimate emission reductions.
Estimate emission reductions. Assess cost-effectiveness.	• Carry out cost-effectiveness analysis with full complement of benefit and cost estimates.
Programme Revision ***(36+ months)***	
Revise or update programme theory hypotheses.	• Revise statement of programme theory based on results of first-round evaluation.
Update estimates of baseline.	• Revise per results of first round.

Categories of evaluation effort and ambition. Comprehensive, accurate evaluations yield many benefits. They provide valuable information concerning the social benefits attained in return for the expenditure of public resources, furnish insights to make future efforts more cost-effective, and direct management attention towards key issues in programme operation. In short, evaluation provides the critical 'content' of the management information system on emission-reduction efforts.

Comprehensiveness and accuracy, of course, come at a cost. Good evaluation practice is not an all-or-nothing proposition. The needs of managers, regulators, and other oversight bodies can often be met without comprehensive evaluations based on extensive original primary data collection. For example, if a programme promotes a technology such as residential compact fluorescent lamps, whose savings have been studied extensively in the field, the evaluation can borrow results for elements of the savings calculations. Of course, such recycling of study results requires careful judgement regarding the comparability of programmes, efficiency technologies, and target markets. Similarly, if the level of programme effort or expected impact is modest (especially in comparison to other efforts in a sponsor's portfolio) then it will make sense to direct evaluation resources elsewhere. More comprehensive efforts should be considered when programmes:

- Receive significant funding (especially in regard to other programme efforts).
- Yield large expected savings, or a large portion of total savings expected from a portfolio of programmes.
- Use a new programme delivery method or promote a new technology.
- Pilot an effort that may be significantly expanded.
- Are scheduled for significant changes in upcoming years.
- Incorporate new or untested programme theories.
- Receive intensive regulatory, legislative, or public scrutiny.

Three ambition levels. The authors of the Danish *Evaluation Guidebook for Energy Saving Actions (SRCI 2003)* have devised a scale of evaluation effort or 'ambition' that is related to the motivations of evaluation sponsors and the rigour of impact evaluation methods deployed. The following descriptions of (levels of) evaluation effort use this structure and adapt it to the evaluation framework laid out above. The authors will refer to these levels of effort in discussing method selection in regard to specific kinds of policy measures. In practical terms, three levels of effort are defined as follows.

Level A: Comprehensive evaluation:

- outcome indicators including net behavioural change,
- impact indicators on energy savings,
- additional internal and external information sources are needed.

Level B: Targeted evaluation:

- including outcome indicators as gross behavioural change,
- some additional information sources.

Level C: Programme review evaluation:

- focus on input and output indicators,
- only use existing (written) information sources.

Figure 1.16 summaries the major characteristics of these three ambition levels, which are presented below in more detail.

Figure 1.16 Overview of Three Ambition Levels of Evaluations

Elements included	**Level A Comprehensive evaluation**	**Level B Targeted evaluation**	**Level C Programme review evaluation**
Quality control of programme tracking data (inputs and outputs)	Review all key data	Review selection (about half) of key data	Verify level of inputs
Analysis of programme tracking data	Ex-ante estimates	Focused on main target(s) of programme	
Estimate of programme costs	Different types of costs	Restricted to programme costs	Verify programme costs
Market characterisation	Broad range of topics	Focused on information for baseline indicators	Indicators for targeted population
Development of programme baseline	Detailed characterisation	Focused on key elements	
Assessment of programme market effects	Combined information on market and baseline information	Focused on elements for gross and net energy savings	Verify level and distribution of outputs
Estimation of gross energy savings	Based on primary data	Estimation on literature or prior evaluation	Use of engineering algorithms
Estimation of net energy savings	Include free-riders and spillovers	Based on general analysis or self-reports	
Estimation of emission reduction	Include local supply conditions	Use of general emission factors	Use of general emission factors

Level A: Comprehensive evaluation. For the purposes of this *Guidebook*, comprehensive evaluations are those that apply the most rigorous cost-justified methods available to characterise programme effects and estimate energy savings specifically attributable to the programme. The results of a comprehensive evaluation should provide the basis for a rigorous cost-benefit analysis, which in turn can be used to assess the value of the programme relative to other investments of social resources.

A comprehensive evaluation will typically include the following elements.

- ***Quality control of programme tracking data (inputs and outputs).*** This step generally includes a review of all key data items collected by the programme for completeness, consistency, and conformity with field definitions. Also, for programmes that involve significant involvement in the development of relatively complex projects, evaluation will typically include ‘end-to-end’ testing of a sample of cases to ensure that project documentation procedures were followed.
- ***Analysis of programme tracking data***. This analysis helps to identify patterns of participation, develop a narrative of programme implementation for use in time series analysis, develop sample plans for customer and vendor surveys and other primary data collection, and develop components of energy impact analyses. The latter will include estimates of total and average *ex-ante* engineering estimates of energy savings by customer segment and measure type. Many saving-estimation procedures involve

making adjustments to these quantities based on the results of field observation, billing analysis, and customer surveys.

- ***Estimate of programme costs***. Most cost-benefit assessment approaches call for comprehensive estimates of programme costs. These include not only the programme's administrative costs, but also the costs of implementing and promoting energy efficiency improvements that were incurred by customers and vendors in the targeted populations.
- ***Market characterisation***. This analysis will generally cover a broad range of topics defined by the technologies and markets addressed by the programme. In terms of the evaluation framework used in this *Guidebook*, the market characterisation serves as the basis for developing the baseline for assessing the net effect of the programme on adopting and promoting energy-efficient products and practices. Topics typically covered in the market characterisation include the following.
  - Market size.
  - Market structure and segmentation.
  - Purchase and technology-adoption decision-making processes by customers and vendors.
  - Barriers and motivations to adopting and/or promoting efficient products and services.
  - Trends in awareness, availability, price, promotion, and market share of efficient products/prevalence of efficient practices.
- ***Development of programme baseline.*** The approach to developing the programme baseline and the indicators used to characterise the baseline will differ depending on the technologies and markets addressed by the programme. For example, the baseline for a new construction programme could be expressed as a set of prescriptive design characteristics or an Energy Utilisation Index (e.g. annual kWh used per square meter for key end uses). However, the baseline for an equipment rebate programme generally includes a baseline market share for efficient models as well as assumptions concerning the annual energy consumption of standard versus efficient models.
- ***Assessment of programme market effects.*** This analysis characterises the effects of the programme on awareness, availability, price, promotion, and market share of efficient products/prevalence of efficient practices. Typically, these analyses use information collected in the market characterisation, combined with one or more baselining techniques to isolate the effects of the programme.
- ***Estimate of gross energy savings.*** A comprehensive evaluation includes estimates of gross energy savings based mostly on primary data concerning changes in energy consumption from a statistically valid sample of targeted facilities or customers. Moreover, the samples must be sufficiently large to yield confidence intervals (around the estimates of total or average gross savings per unit) that satisfy regulatory requirements. The methods applied to develop these estimates may include any of the approaches discussed above or, more likely, perhaps a combination of these methods.
- ***Estimate of net energy savings.*** A comprehensive evaluation will include estimates of net energy savings that account, either implicitly or explicitly, for the effects of free-ridership and spillover. These estimates will usually employ cross-sectional or historical analyses, often in combination with each other or with other approaches discussed above.

- ***Estimate of emissions reductions.*** A comprehensive evaluation includes an estimate of emissions reductions that reflects local energy supply conditions.

Level B: Targeted evaluation. Targeted evaluations are also designed to yield relatively rigorous estimates of programme energy savings and emission reductions. However, due to schedule and budget constraints it may not be possible to do a thorough job on all the evaluation elements identified for the Level A evaluation. This is the case for most evaluations that are carried out. Level B evaluations generally contain the following elements.

- Quality control of programme tracking data (inputs and outputs).
- Analysis of programme tracking data.
- Estimate of programme costs.
- Market characterisation. In a Level B evaluation, market characterisation, if it is performed, generally focuses closely on gathering information needed to formulate baseline indicators.
- Assessment of Programme Market Effects. In a Level B evaluation, assessing market effects generally focuses on elements required to estimate gross and net energy savings.
- Estimate of gross energy savings. In a Level B evaluation, information from secondary literature or prior evaluations may be used to estimate one or more elements of the gross savings estimate. Of course, this approach requires the application and documentation of various assumptions concerning the similarity between the current programme domain and the sources from which other information was taken.
- Estimate of net energy savings. In a Level B evaluation, estimates of net energy savings may rely on the results of only one general analytic approach – very often on self-reports from programme participants.
- Estimate of emissions reductions.

Level C: Programme review evaluation. This is the least intensive form of evaluation. Its primary objectives are to:

- Verify the level and distribution of programme outputs such as number of facilities assisted, number and type of equipment units rebated, numbers of trainees at seminars, etc.
- Verify the level of inputs to the programme – fiscal expenditures, materials, etc. – to support late cost-benefit analysis.
- Characterise actual changes in key indicators for the targeted population of facilities or customers that occurred during the programme period. These could include changes in levels of customer awareness of energy-efficient products, in the market share or the adoption pace of those products, in the percentage of efficient products that retailers display etc. It could also include estimates of energy savings based on the application of engineering algorithms to track system data on the number and type of programme participants or purchases/installations subsidised.

Evaluations at this level essentially serve to check the progress of programme implementation and customer response in relation to programme plans and assumptions. Generally they will include neither extensive primary data collection to estimate energy savings nor rigorous estimation of baselines or net savings.

In general, a comprehensive evaluation is conducted for major, longer timeframe policies, while smaller measures are only reviewed. Figure 1.17 shows the level of ambition, defined after the evaluation has been completed. This can help when discussing the level of ambition at the start of an evaluation. In the USA a large number of utility programmes were evaluated during the 1990s on a detailed level, i.e. effort levels A and B.

Figure 1.17: Evaluation Case Examples and Level of Ambition

Case examples	Country	Level A Comprehensive	Level B Targeted	Level C Review
Policy type Regulation				
Building codes	Belgium		x	x
Energy Efficiency Regulations for Residential Equipment	Canada		x	
Energy management scheme for large buildings	Denmark		x	
Minimum energy performance standards	Korea			x
Energy Performance Standard (EPS) for houses	Netherlands	x	x	
Policy type Information				
Local energy efficiency information centres	Belgium			x
Energuide for houses	Canada		x	
Energy labelling of small buildings	Denmark		x	
Free-of-charge electricity audit	Denmark	x		
Project ‘Red-Hot’ (element of stand-by campaign)	Denmark	x		
The ‘A’ campaign 1999	Denmark		x	x
Promotion campaign for efficient ventilation	Denmark	x		
Information campaign (2001)	France		x	
Local energy information centres (Espaces Info Energie, EIE)	France	x	x	
Audits (“Aides a la decision”)	France		x	
Energy audits in industry	Korea			x
Energy audits in buildings	Korea			x
Energy Efficiency Rating Labelling	Korea	x		
Information centres in local region	Sweden	x	x	
Information and education programme 1998-2002	Sweden		x	
Policy type Economic				
Criteria adopted for the evaluation of primary energy savings in end-uses/ EE Certificates	Italy			x
Rebate programme for highly efficient electric inverters	Korea			x
Financial incentives for DSM	Korea			x
Energy premium scheme households	Netherlands	x	x	
Energy Investment Reduction (EIA and EINP)	Netherlands			x
Policy type Voluntary Agreements				
Canadian Industry Program for Energy Conservation	Canada	x	x	
Voluntary Agreements	Korea			x
Voluntary Agreements on Industrial energy Conservation 1990 - 2000	Netherlands	x	x	
Eco-energy	Sweden	x	x	
Combined policy Measures				
Rebate programme for household appliances	Belgium	x	x	
STEM programmes	Sweden	x	x	

Evaluation costs. The evaluation costs will vary according to the size of the programme and its ambition levels. There are indications that the monitoring and evaluation activities range up to 8-10% of the programme budget[17]. In general the evaluation budget is (much) smaller. The team of experts has insufficient facts to back up cost level indications for evaluation costs relating to programme costs in general and, more precisely, in combination with ambitions.

[17] Guideline for the Monitoring, Evaluation, Reporting etc. 1999, page 72

2. EVALUATION OF REGULATION POLICY MEASURES AND PROGRAMMES

2.1 Introduction

The regulation policy measures and programmes addressed in this guidebook include building codes and performance standards for (in-house) equipment. Case examples for building codes from Belgium, Denmark, and the Netherlands are used to illustrate the theory (see Table 2.1). A detailed description of the cases can be found in Volume II of the *Guidebook*. A case example on the Swedish building codes is not included as this has not yet been evaluated.

Table 2.1 Subcategories and case examples for regulation policy measures and programmes

Subcategories	Case examples	Country
Building Codes and Enforcement	Building codes	Belgium
	Energy management scheme for large buildings	Denmark
	Energy Performance Standard (EPS) for houses	Netherlands
Minimum equipment energy performance standards	Energy Efficiency Regulation for Residential Equipment	Canada
	Minimum energy performance standards	Korea

The following section defines the two subcategories of regulation measures and states their main objectives. Sections 2.3 and 2.4 then analyse each of the two subcategories according to the seven key elements discussed in Chapter 1:

Element 1: Policy measure theory used
Element 2: Specification of indicators for the success of a measure
Element 3: Development of baselines for the selected indicators
Element 4: Assessment of outputs and outcomes
Element 5: Assessment of energy savings and emissions reductions and other relevant impacts
Element 6: Assessment of cost, cost-efficiency and cost-effectiveness
Element 7: The level of evaluation effort

Based on the country examples and the discussions at the experts meeting, we see the following main issues taken into account for evaluating regulation policy measures and programmes:

- Baseline:
  - The baseline (including a good description of the assumptions) should be developed prior to the implementation.
  - The handling of free-riders should be well documented.
  - The impact of an update of codes (especially if and how much more ambitious the code is than formalised in normal practice) should be specified.
- The use of value from test procedures and/or the real values in practice should be included in the assessment of energy savings.
- In the assessment of outcomes attention should be given to measuring the compliance with regulations and a system of (independent) control.

2.2 Objectives and Main Types of Regulation Policy Measures

In this *Guidebook*, the term 'Regulation' refers to laws and implementing rules regarding requirements for devices to advance energy-efficient design and construction. These requirements generally apply to new or renovated buildings or to (in-house) equipment and appliances. The two main subcategories of regulation policy measures for energy efficiency include building codes and minimum performance standards. Each of these subcategories will be discussed in separate subsections below.

2.2.1 Building Codes

Building codes specify how buildings (or subsystems of buildings) must be constructed or perform. Most codes apply to both residential and non-residential (commercial) buildings, although the exact requirements usually differ for the various categories of buildings. The codes are generally mandatory and enforceable, although some countries, such as Canada, still continue to use voluntary codes. Generally, the codes affect the energy demand for heating or cooling of a building, although some measures also incorporate electricity demand for ventilation etc. The World Energy Council[18] distinguishes five types of (thermal) building codes:

Type 1: Envelope component approach / Component standard

This was popular in early European building codes in the 1970s, but is still regularly used for single-family houses. It considers the heat transfer (heat losses or gains) through individual components of the building shell, such as external walls, roof, windows, etc.

Type 2: Overall envelope approach

Type 2 building codes are more flexible than Type 1, as this approach only sets a limit on the overall heat transfer through the building envelope, thus allowing reduced heat transfer through a single part of the shell (e.g., walls, roof, windows) to compensate for more heat transfer through another part.

Type 3: Limitation of heating/cooling demand

In addition to the transmissions through the building envelope, this type of code also includes the contributions from ventilation losses or gains, passive solar gains through building components (in particular through windows), and internal heat sources.

Type 4: Energy performance standard

The type 4 building code is the first to include an integrated approach: it considers the whole building as a system. It integrates not only the demand for heating and cooling, but in addition all (or most) of the building equipment, such as heating and air-conditioning systems, energy for ventilation, hot water preparation, pumps, lifts, etc. In particular, it also includes all active solar energy gains from solar collectors, photovoltaic units, etc. The new EU Building Code is an example of this type of code.

18 See WEC 2001, pp. 57-58.

Type 5: Life-cycle standard

This type is still under research; no country has yet realised such a standard. In addition to the items covered in Type 4, the life-cycle standard also covers the energy used to produce the building materials. This issue becomes increasingly important when the direct energy consumption of buildings is reduced.

According to the World Energy Council [WEC 2001, pp. 57-59], the building codes currently prevailing are Type 3 and Type 4, which are performance-based (as opposed to component-based approaches). However, the World Energy Council [WEC 2001, p. 186] states that there is now a strong tendency towards the integrated approach of Type 4, which give designers more flexibility to meet energy-reduction standards in the most cost-effective way. This integrated approach not only includes the quality of insulation in the building, but can also take into account heating and cooling installations, energy for ventilation, lighting installations, position and orientation of the building, heat recovery, active solar gains and other renewable energy sources.

The performance-based approaches can be complemented with standards on specific building materials (e.g. insulation, windows) or equipment (e.g. boilers), in order to ensure the dissemination of the most efficient equipment in the retrofitting of existing buildings. Standards on the energy performance of (in-house) equipment will be discussed in the next section.

2.2.2 Energy Performance Standards for Equipment

There are three types of efficiency standards[19]: prescriptive standards, performance standards (MEPS), and fleet-average standards. Prescriptive standards require that a particular feature or device is installed in all new products (such as catalytic converters in cars). Performance standards prescribe a minimum efficiency (or maximum energy consumption) that manufacturers must achieve in each product, and the class-average standards specify the average efficiency of a manufactured product. The prescriptive standard can be used for the market, in combination with each of the other two standards, and each standard can be either mandatory or voluntary.

This chapter addresses the mandatory performance standard[20], the most commonly applied standard for household and office appliances. Prescriptive standards are often part of building code requirements (see § 2.3), and are less flexible than performance standards, as the latter do not specify the technology or design details of a product, only the minimum energy performance. Fleet-average standards are more flexible than performance standards, as they allow each manufacturer to select the level of efficiency for each model, as long as the overall average is achieved. However, the standard requires an elaborate and sophisticated procedure for assessing and enforcing compliance and adds considerable complexity to the manufacturer's production and shipment schedules. These types of standards are only used in Switzerland and Japan.

---

[19] See CLASP: Energy-Efficiency Labels and Standards: A Guidebook for appliances, equipment and lighting, by Wiel and McMahon, (pp. 7-9).

[20] Note that we define the voluntary standards as a form of information supply.

Performance standards, usually known as minimum energy performance standards (MEPS), are well-known for refrigerators and washing machines, but can also apply to all kinds of energy-using devices such as (household) appliances, office equipment, transformers, electric motors, and packaged heating, ventilating, and air conditioning (HVAC) equipment. Most standards concern electrical appliances, although standards for gas appliances, for example, also exist. Performance standards are often used in combination with labelling to provide consumer information on the energy consumption of appliances (see also Chapter 3, which covers information as a policy measure).

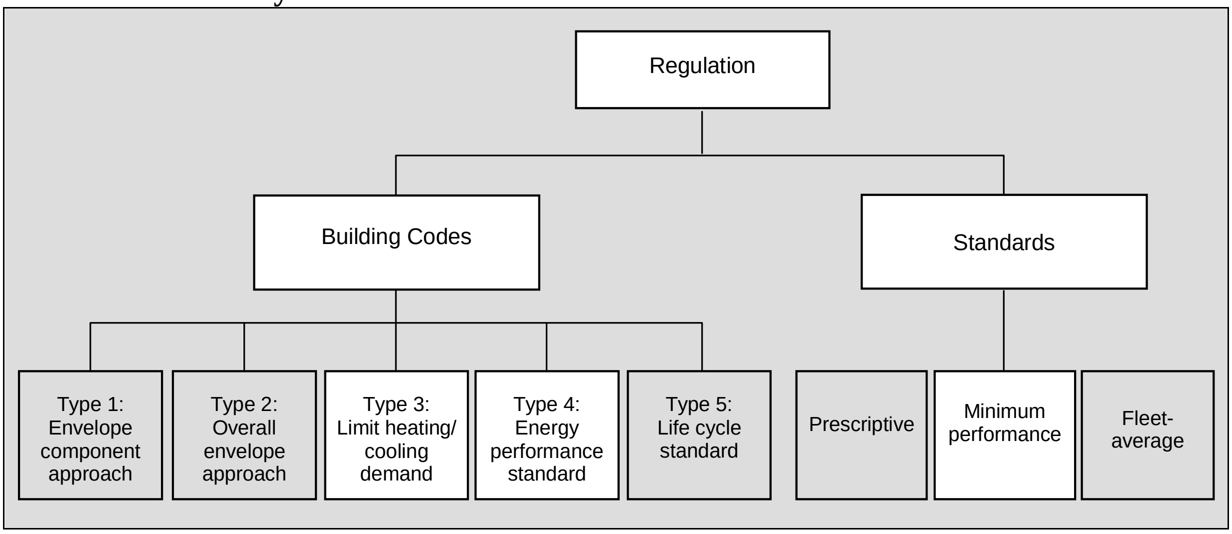
As stated above, the mandatory performance standard imposes a minimum energy efficiency rating or maximum energy consumption for all products on the market, and prohibits the sale of equipment that is less energy efficient than the minimum level. The standard usually includes a well-defined protocol or test procedure that prescribes how to measure and rank the energy efficiency of a particular product. The remainder of this analysis focuses on standards for (electrical) appliances for households and offices only, as industrial appliances are a deviating category and building equipment can largely be covered under building codes.

Countries use various ways to set levels for efficiency standards[21]. For example, Europe uses a statistical approach, where the energy efficiencies of appliances already on the market serve as a basis for setting the minimum level. The standard is set at a level sufficient to obtain a 10-15% improvement in the average energy efficiency of appliances on the market. Other countries (such as the USA) base their standards on cost-benefit evaluations, where a fixed number of years for the return on investment determine the energy efficiency level of appliances.

2.2.3 Overview of Regulation Policy Measures and Subtypes

Figure 2.1. shows the two subcategories and underlying types of regulation policy measures. In the remainder of this *Guidebook*, only the performance-based building codes (i.e. Types 3 and 4) are included in the analysis, and only the performance standards.

Figure 2.1 Categories and subtypes of regulation policy measures: the white boxes are included in the analysis



21 See World Energy Council (WEC 2001, pp. 70-71).

Sections 2.3 and 2.4 analyse each of the two subcategories according to the seven key elements discussed in Chapter 1:

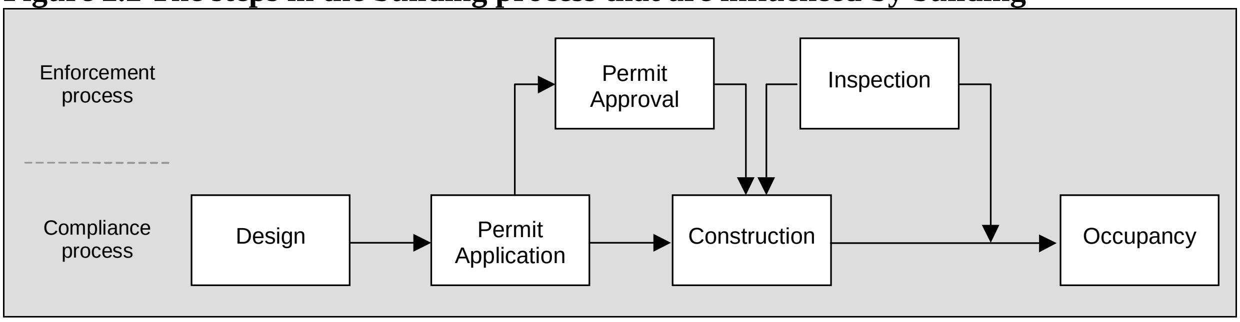
(i) Policy measure theory used.
(ii) Specification of indicators for the success of a measure.
(iii) Developing baselines for the selected indicators.
(iv) Assessment of outputs and outcomes.
(v) Assessment of energy savings and emissions reductions and other relevant impacts.
(vi) Assessment of cost, cost-efficiency and cost-effectiveness.
(vii) The level of evaluation effort.

2.3 Building Codes Policy Measures and Programmes

2.3.1 Policy Measure Theory

As discussed in Chapter 1 (see section 1.3.1), the theory behind a policy measure provides the basic framework for evaluating that measure, and the theory should address at least the domain and the effects hypotheses of the policy measure. However, in order to determine the domain and effects hypotheses, it is useful to identify the steps in the building process that are influenced by a building code. These steps are shown in Figure 2.2, which is used to discuss the domain and effects hypotheses for building codes in the sections below.

Figure 2.2 The steps in the building process that are influenced by building

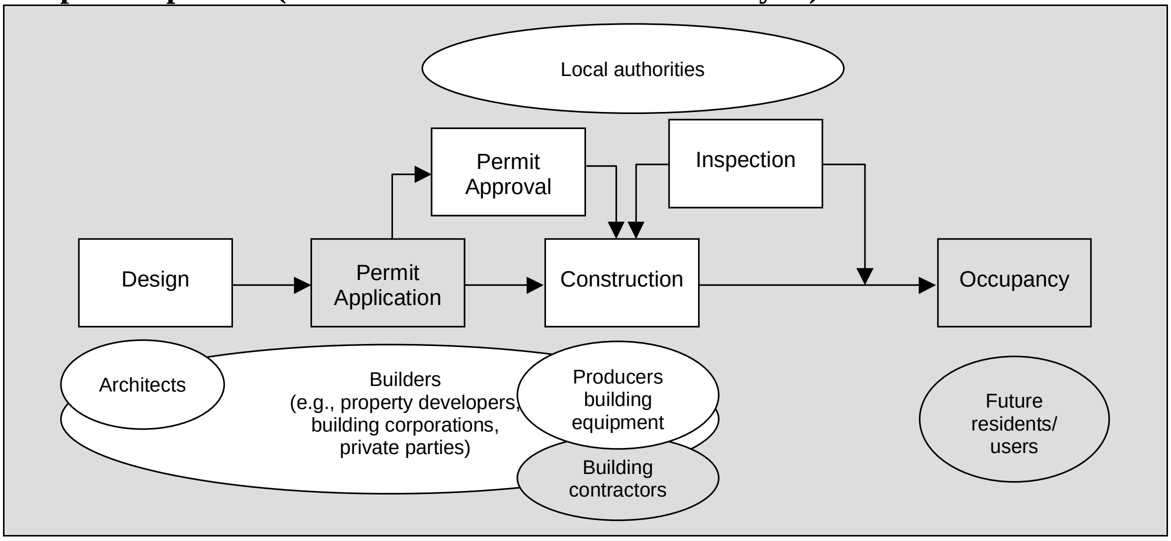


codes (Source: Xenergy 2001, p.s1)

Policy Measure Domain

The stakeholders that are (likely to be) affected by the building process are shown in Figure 2.3. However, not all stakeholders will be included in the domain of the building code: the domain depends on which steps of the building process are taken into account. Most programmes concerning building codes focus on the design and construction phase, where the architects, builders (i.e. stakeholders that assign projects to building contractors, such as building corporations and property developers –but also private builders), and the producers of building materials and equipment are the main target groups. However, local authorities can also largely influence the effects of building codes by the manner in which they enforce the code ex-ante (through permit approval) and ex-post (through inspection at the building site). So the remainder of this analysis focuses on these groups as the main stakeholders in the building code domain. Nonetheless, other stakeholders, such as the future residents/users can – through their behaviour – also influence the actual energy performance of a building significantly. The hypotheses on effects for the relevant stakeholders are discussed in the next section.

Figure 2.3. Stakeholders that are likely to be affected during the building and compliance process (white cells are included in the analysis)

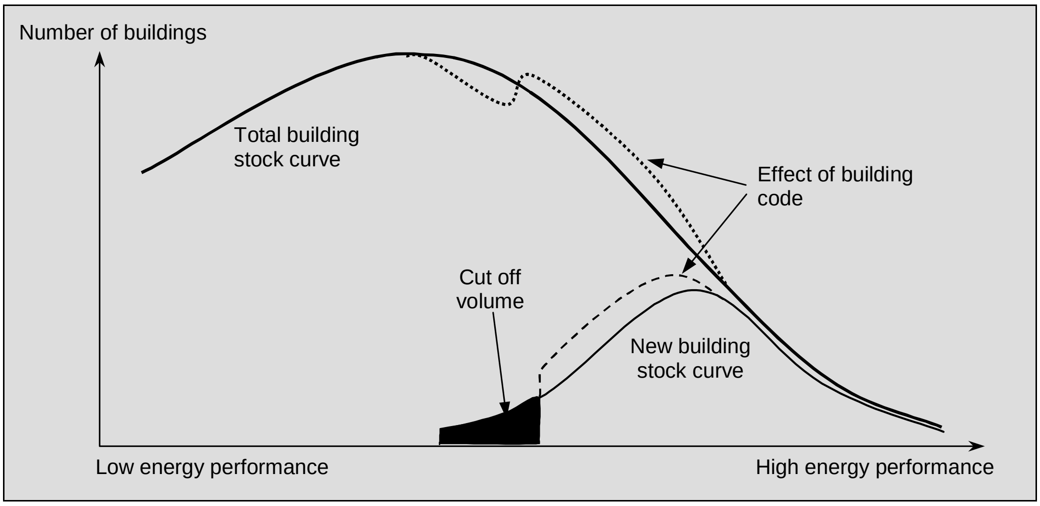


Policy Measure Hypotheses

Two types of building codes are addressed in the analysis: Type 3 (Limitation of heating/ cooling demand of buildings, but excluding building equipment); and Type 4 (an integrated energy performance standard for energy consumption of the building system). However, since the domains and hypotheses of these two types are largely the same, we will not address them separately and only point out where they differ from each other.

The theory underlying the building codes assumes that architects, builders, producers of building equipment, and contractors will (by default) apply energy-saving measures in their designs and constructions. In turn, the local authorities will check whether the designs and actual constructions comply with the code. The code must ensure that the buildings with the worst energy performance can no longer be constructed, but this only applies to the new-to-build buildings, not existing buildings, as shown in Figure 2.4.

Figure 2.4 The theory underlying building codes: the new-to-build buildings with the worst performance are cut off from the market, but the effect on the total building stock is small



Usually, the government also announces that the code will become more stringent in the future, in order to encourage the construction of buildings that perform a lot better than the code requires, as architects and builders will anticipate the future requirements. This causes the new building stock curve in Figure 2.4 to shift slightly to the right, and therefore the curve of the total building stock shifts to the right, although the latter effect will be marginal. Table 2.2 gives an overview of the assumed effects for each of the domains.

Table 2.2 Effects hypotheses on regulation policy measures for various stakeholders in the domain of building codes

Building Codes		Domain specification		Hypotheses on outcomes		Hypotheses on desired impacts
Type 3, Type 4	→	**Architects** *Hypotheses on:* Awareness of building code, tools, and future tightening up	→	Energy-efficient building designs Anticipation on future more stringent code levels	→	Type 3: Reduced energy consumption per building for heating and cooling purposes Type 4: Reduced energy consumption per building
		↑↓		↑↓		
Type 3, Type 4	→	**Builders** *Hypotheses on:* Awareness of building code, tools, and future tightening up Enforcement of code ex-ante (permit approval) Enforcement of code ex-post (inspection)	→	Choice for more energy efficient building designs (anticipation) Increased frequency of using code-prescribed building methods. Building codes for new building also influence the existing buildings through retrofit	→	Type 3: Reduced energy consumption per building for heating and cooling purposes Type 4: Reduced energy consumption per building
		↑↓		↑↓		
Type 3, Type 4	→	**Local Authorities** *Hypotheses on:* Knowledge of building code Enforcement of codes ex-ante (permit approval) Enforcement of codes ex-post (inspection)	→	Work approach is adjusted to building code Applications for building permits are correct and comply with code Actual construction of buildings is consistent with the permit specifications	→	Type 3: Reduced energy consumption per building for heating and cooling purposes Type 4: Reduced energy consumption per building
Type 4	→	**Producers of Building Equipment** *Hypotheses on:* Awareness of building code, tools, and future tightening up	→	Energy efficiency of equipment for heating demand in buildings increases	→	Type 4: Reduced energy consumption per building

2.3.2 Specification of Indicators

Indicators provide information on the effects of a policy measure, and are used to measure the outcomes and impacts. Based on the effect hypotheses (discussed in the previous section) the indicators for building codes can be divided into the following categories:

- Level of awareness and knowledge.
- Level of adoption of practices (learning by doing).
- Design of buildings.
- Level of enforcement and compliance.
- Changes in energy consumption of buildings.

Chapter 3 addresses the indicators level of awareness and knowledge, as this mainly concerns information and communication activities. Examples of possible indicators associated with the other categories are listed in Tabele 2.3 below.

Table 2.3 Examples of outcome and impact indicators for building codes

Categories of Indicators	**Example indicators**
Level of adoption of practices (learning)	• % of architects that apply energy-efficient techniques and constructions in their design • % of builders that adopt energy-efficient techniques during construction • % of producers that produce energy-efficient building equipment
Design of buildings	• % of passive solar designed buildings • % of rejected building permit applications
Level of enforcement and compliance	• % of buildings that comply to all building code requirements
Energy efficiency of building equipment	• Heat transfer coefficient (k-value or U-value, in W/m^2K) or r-value (Km^2/W)) of components • Energy efficiency of equipment
Changes in energy consumption of buildings	• Changed energy consumption per building for heating and cooling purposes (W/m^2 or kWh/m^2 per year or kWh/m^3 per year) • Changed energy consumption per building ($kWh/m^2/yr$ or $kWh/m^3/yr$)

The approaches to determine the indicator scores vary per category and per indicator. The actual application of energy-efficient techniques in design and construction can be achieved through interviews and questionnaires, but can also be assessed by analysing building plans and permits, inspecting buildings under construction, and inspecting fully constructed buildings. The number of buildings that comply with the building code in a country can generally be deducted from the data held by the (local) authorities that enforce the building code. However, ex-post evaluations of code compliance are generally rare.

The energy efficiency of building materials and equipment is commonly specified by the k-value or R-value of the components, or by the energy efficiency of the equipment. The changed energy consumption per building can be measured with the mean annual heating/ cooling demand in kWh/m^2 or kWh/m^3 (Type 3) and the mean annual energy consumption (kWh/m^2 or kWh/m^3) of a building. This is calculated with the help of computer models or special software using the energy consumption of standardised 'reference buildings', but in some cases real practice measurements have been taken[22].

[22] For example, City of Fort Collins (2002) and Arkansas Energy Office (1999).

2.3.3 Development of Baselines

A reference situation needs to be determined in order to assess the effects of building codes. This 'baseline' should state the situation in the absence of the building code, particularly the average energy performance of the new buildings. The baseline should preferably be determined ex-ante, but it is often defined ex-post. To determine the baseline ex-ante you can use, for example, surveys among designers, design specifications in approved permits, on-site audits, and simulation. Scenarios are also often used to develop baselines. Scenarios are helpful when uncertain factors that influence the desired impact of a policy measure need to be incorporated into the analysis. The introduction of the new building code results in a different scenario, so that the difference between the two scenarios clearly reflects the impact of the building code.

The ex-post baseline for a new code is relatively easy to determine if building codes have already been enacted in the past: the baseline is simply the situation in which the new buildings would have complied with the previous code (including a certain factor for non-compliance). Otherwise, the baseline consists of new buildings that would have been built with the average energy performance prevailing just before the introduction of the new code. Note that the baseline generally only addresses new-to-build buildings and not the already existing stock, as the effects on the total stock will be quite small, and therefore more difficult to determine.

2.3.4 Assessment of Output and Outcome

The *output* of building code regulation is, of course, the code itself (legislation), but can also include tools for enforcement, awareness programmes (such as building code workshops and training courses) or information material (such as guidebooks, reports, brochures and leaflets). For example, the Dutch government, when introducing an energy performance standard for new buildings, also introduced a new tool known as 'EPCheck' to facilitate permit approvals as well as building site inspections.

However, determining the *outcome* of building code regulation is not as straightforward. Most of the evaluations use the (approved) building permits to assess the outcome of building codes[23]. These evaluations are generally implemented through computer modelling, using theoretical 'reference buildings' with standardised energy performances to compare the energy performances of approved building designs. A complicating factor is the timeframe that exists between permit application and actual construction, which can be quite substantial. Permit applications that are made before the code was enacted do not have to comply with the code, even if the actual construction of the buildings takes place well after its introduction. This creates a time lag between the introduction of the code and the actual effect of the code in terms of the energy performance of buildings, and this time lag should be accounted for in the evaluations.

Another complicating factor is the fact that new buildings are not always constructed strictly according to the building permits, even if the local authorities manage to enforce the code during permit approval. Due to limited resources, building site inspection is often not a first

---

[23] Examples of ex-post evaluations are included in City of Fort Collins, (2002), Arkansas Energy Office (1999), Australian Greenhouse Office (2000), and WEC (2001).

priority for local authorities, which results in non-compliance with the code. A clear example of this is the extensive evaluation of the building code in the City of Fort Collins in 2002. This evaluation revealed that the energy performances actually measured as a result of a new building code (Type 1) were *less than half* the performances anticipated through models. Evaluators attribute this discrepancy to five factors, one of which was code non-compliance (the model assumed full code compliance)[24]. They conclude that non-compliance is largely due to a rather low effectiveness of quality control procedures. Similar findings for non-compliance were reported during a study in Massachusetts, concerning a Type 3 (performance-based) building code and a study in Washington State[25].

However, if adequate attention is paid to enforcement (e.g. through tools, courses, handbooks, etc.), one of the outcomes of a building code can be that the local authorities adopt a new work approach whereby the enforcement of building codes is improved. So ex-post evaluations of the energy performance of buildings are rare and little information is available on the percentage of constructed buildings that actually comply with the required code. One of the known exceptions is the two-year RER (residential new construction) study on newly constructed low-rise residential homes across California. (RER, 2001). The study found that in total 57% of these houses were compliant (52%) or overly compliant (5%), while 14% were non-compliant and 29% were indeterminate (RER, 2002 main study page EX-11).

In a sample of 200 newly built houses in Flanders, built both before and after the legal introduction of the building codes, many characteristics were measured (dimensions and type of rooms, glazing, walls, etc.; insulation thickness; indoor temperatures; energy consumption). This data allowed the calculation of the average insulation value (including the effect of thermal bridges), and compliance with the building code could be tested. To address this issue, officials from the Flemish Administration of Energy (ANRE) increased the number of on-site checks, on top of the existing administrative checks. In 1999 they carried out 742 on-site random checks of residential building sites, as opposed to the previous 450 per year. In the years 2000 and 2001, these checks dropped back to 622 and 540 respectively. The administrative checks involve a conformity control of every insulation form that is submitted with ANRE. These insulation forms are obligatory for obtaining a building permit. The on-site checks take place on the basis of these insulation forms. During the inspections, the global K-value is not verified, but rather the individual U-values of the wall elements. It seems that (considerable) differences continue between the information on the insulation form and reality. The problem is that the officials can only warn architects and principals in the case of non-compliance, but cannot impose sanctions under the present insulation decree (see the case example for Belgium in Volume II).

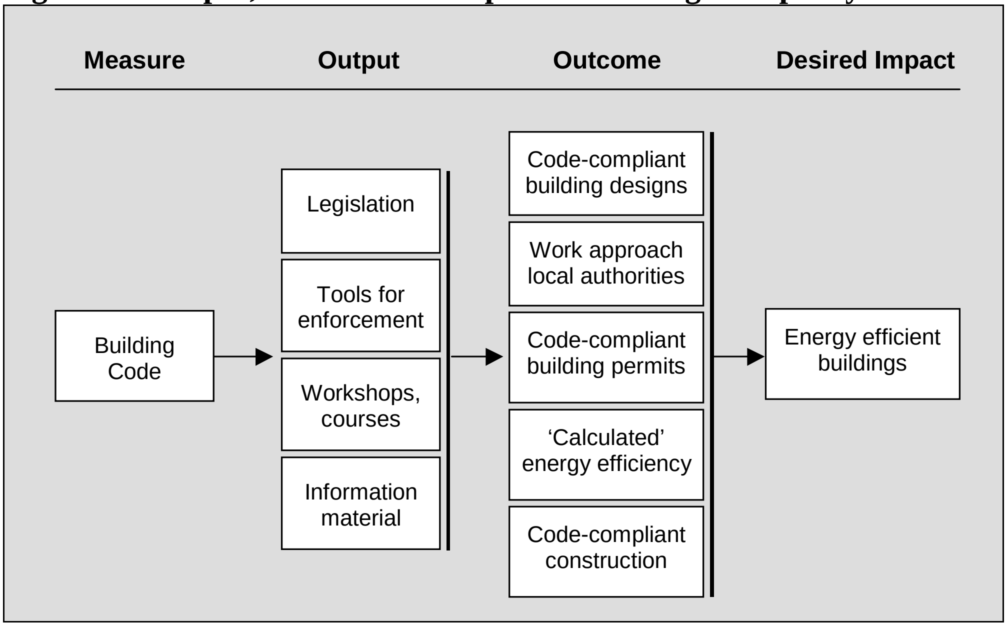
Energy bills can be used to obtain more information on the actual energy performance of buildings, but these bills also reflect the behaviour of the building's residents/users, which makes it difficult to compare results with baseline situations. In fact, the behaviour of the building's future residents/users has a major influence on the *impact* of building code regulation (i.e. the actual energy performance of the buildings). So the non-compliant building designs, lack of adequate enforcement, the non-compliant construction of buildings, and the behaviour of users all account for the fact that the energy performance of newly constructed buildings is generally less than that anticipated using computer models (see also Section 2.3.5

[24] Other factors were mainly related to assumptions on temperatures (e.g., thermostat setpoint, basement temperature, internal gains).

[25] See Xenergy (2001, 5-3), and WSU (1997, p. 7) respectively.

on energy savings). Figure 2.5 gives an overview of the (possible) output, outcome, and impact of building code regulation.

Figure 2.5 Output, outcome and impact of building code policy measures



Thus, with respect to the energy performance of buildings, only if the building designs are code compliant, if the code is strictly enforced, *and* if building users show 'average' behaviour in energy use, is the actual impact of the code equal to the calculated outcome. The energy savings and emissions reductions that are associated with the building code regulation are addressed in the next section.

2.3.5 Assessment of Energy Savings and Emissions Reduction

Analysis of energy savings

The energy savings that result from building code regulation are determined by comparing the effects of the building code with the baseline situation. As explained in the previous section, it is difficult to determine the *impact* of code regulation. Relatively few countries have carried out field evaluations of their building codes to assess the real performance of buildings. The energy consumption per building would be the main indicator, but this covers all aspects of the building in use, not just the energy performance of the building itself. This includes the efficiency of electrical appliances, the climate or weather conditions, the behaviour of building users, as well as energy costs etc. Therefore, the energy savings are usually determined through the output, using the energy performance of building designs described in building permits. The energy savings can be derived using the number of new building designs that comply with the new code, but that would have been built with an average energy performance in the absence of the code. The more stringent the energy code, the higher the energy savings. However, as previously mentioned, non-compliance with the code may result in too optimistic values for energy savings calculated with models.

In addition to non-compliance, there are two other reasons that make it difficult to assess actual impacts of building codes[26]. One of these reasons is the slow penetration of efficient (new) buildings into the total stock of buildings, resulting from the vast amount of already existing buildings and the long lifetime of these buildings. The effects of building codes will thus only become apparent in the long term, when the energy-efficient buildings encompass a significant share of the total stock. This shows the importance of also considering measures for the existing building stock.

A factor that further complicates the evaluation of building codes concerns the recurrent revisions of the codes over a period of time, which makes it difficult to attribute the effects to a specific period. Another important factor that offsets the effects of building codes is the increasing demand for comfort (i.e. larger buildings). For instance, in the Netherlands, a large building can have the same performance as a small building (applying the same energy measures results in the same energy performance, regardless of the size of the building, see Box 1). The large building will obviously have larger surfaces that need to be cooled or heated, which leads to extra energy consumption, even though the energy performance is the same as for the small building.

An issue that is closely linked to the demand for comfort is the rebound effect: sometimes, the energy savings are offset due to the fact that users act differently as a result of these savings. For example, people tend to leave the lights on if their house is equipped with energy-efficient lighting, because ‘it is much nicer with the light on, and it is cheap’. Or people use significantly more hot water in buildings with solar collectors, because ‘the hot water is free’.

Box 1 Building code regulation in the Netherlands

In the Netherlands, the energy performance standard for buildings is a measurement of the primary energy consumption related to space heating, ventilation, water heating, and lighting (including the energy for pumps etc.). Energy consumption is based on standardised consumer behaviour and standard conditions. The calculation of the energy performance is achieved through the energy performance coefficient (EPC), which is formulated in such a way that buildings with the same energy measures have the same energy performance (larger buildings can thus have the same performance as small buildings). The EPC is given as the following formula:

$$EPC = \frac{Q_{tot}}{(65 \cdot A_{shell}) + (330 \cdot A_{use})}$$

Where Q_{tot} Total primary energy consumption of the building
A_{shell} Surface of the building shell
A_{use} Surface of floors in heated zones

The standard includes criteria for determining the total primary energy consumption. The lower the EPC value, the better the energy performance of that building. The energy performance standard prescribes a maximum value for the EPC. The standard has become more stringent over the years. For homes[27], the EPC started at 1.65 in 1992, changed to 1.4 in 1996, and was reduced to 1.2 in 1998. Since 2000, the Dutch government has set the residential standard at 1.0, which corresponds to an anticipated natural gas consumption of 1000 m^3. As of 2006, this standard will become even more stringent (maximum EPC of 0.8 for residences).

Ecofys (2004, p. 38) has estimated that the energy performance standard has resulted in cumulative energy savings of around 2.7 PJ of primary energy within the period 1995-2002, which corresponds to 0.15 Mton of avoided CO_2 emissions.

26 See WEC (2001, p. 65) for an overview of factors that complicate the analysis of energy savings.

27 Different standards for EPC values are set for different building types.

Despite the non-compliance issue and user behaviour, positive effects of building codes are reported in the literature. For example, an ex-post evaluation in Massachusetts reports energy savings of 23.4% relative to the baseline[28]. The Fort Collins study showed that annual natural gas requirements fell 16% (approximately 18.5 GJ per home) as a result of the code. As the code was a Type 1 building code, with little focus on cooling requirements, the savings occurred mainly for space heating and water heating (if this was done with natural gas appliances), while leaving electricity consumption unchanged. So inadequate compliance or enforcement can substantially diminish the benefits of new building codes.

The World Energy Council reports that cumulative energy savings for new residences in the European Union amount to around 60% (on average) compared to residences built before the first oil crisis. However, this figure does not reflect the effect of a single code, but of several codes throughout the years, with different requirements per country. In the near future, the introduction of the energy performance for buildings directive (EPBD) will standardise the use of building code regulations throughout the EU.

Analysis of emissions reduction

The analysis of emissions reductions is based on the avoided use of particular fuels. In the case of building codes the demand for space heating and hot water is usually met using natural gas and, in some cases, electricity or oil. To calculate the reduction in emissions, the energy savings are multiplied by the emissions factors or coefficients that are attributed to each type of fuel, as explained in Chapter 1. The focus is usually on CO_2 emissions, including other greenhouse gases that are converted to CO_2 equivalent emissions.

If the energy savings consist of a mix of fuel types (i.e. in the case of electricity), the emissions factor of the mix can be determined through the percentages of each fuel type. In the Netherlands, for example, the average fuel mix for electricity in 2002 consisted of the fuels listed in Table 2.4.

Table 2.4 Average fuel shares for electricity production in the Netherlands in 2002

Fuel Type	**Share in Dutch electricity mix**	**g CO_2 eq/MJ**
Natural Gas	58.2%	56.8
Coal	31.3%	106.4
Nuclear	4%	0
Hydro	0.1%	0
Wind	1.3%	0
Other Renewable	5.1%	0
Total Dutch Electricity	**100%**	**66.4**

NB: Note that no CO_2 emissions are associated with electricity production from nuclear power, hydropower and wind energy. Also note that the emission coefficients refer to primary energy consumption, and conversion efficiencies from primary energy to end-use energy should be accounted for in the calculations.

28 See Xenergy (2001, p. 5-3), City of Fort Collins (2002, pp. 24-26), and WEC (2001, p. 59).

The CO_2 emissions from electricity can also be determined using a top-down approach: total CO_2 emissions from electricity production are divided by the total energy consumption used for electricity production.

2.3.6 Calculation of Costs, Cost-efficiency and Cost-effectiveness

Total costs include the costs of developing the new building code, costs associated with introducing the code (including development of tools, course material, workshops, handbooks, training, etc.), the costs of enforcing the code, and the costs of evaluating the code. There is little data available on these costs, but in the literature, building codes are reported to be very cost-effective policy measures[29]. Some reports mention costs. For example, in the Netherlands, Ecofys estimates that the costs of implementing the Dutch building code for residences (i.e. the *Energy Performance Standard*) over the period 1995-2000 amounted to 4-7 million euro, while the costs for enforcing the code are estimated to be 1-10 million euro. Estimates of the overall costs for the government range from 6-16 million euro (these estimates have a relatively large range due to the many uncertainties in the underlying data). In terms of cost-effectiveness, Ecofys mentions 0.2-0.8 euro for every GJ of energy saved, which corresponds to 4-14 euro per ton of avoided CO_2 emissions.

WSU (Washington State University) reports that in 1991, checks on building plans and on-site audits of buildings each cost an average of $500, as part of a compliance study on non-residential buildings. However, there was a considerable variation in costs due the range of size and complexity of buildings. Simulation modelling of energy performance was said to amount to around $25,000. The costs of enforcement are often underestimated. The City of Ford Collins states that the 1996 code changes in the city represented a significant increase in workload for the building department, which was already dealing with under-capacity. This led to confusion and inconsistent enforcement and compliance. Although these problems were overcome, the early deficiencies have had lasting effects, which suggest that future code changes should be implemented only when there are sufficient resources available to effectively support the changes.

More data are available concerning the costs of building codes for the investors/end-users. The WEC mentions that the additional costs of buildings are usually limited to a few percentage points, as many countries limit their building code to the economic potential of energy efficiency in buildings, in order to keep the additional costs to a minimum. Although the measures taken to comply with a building code induce additional investment costs, they will (in principle) result in energy savings, and thus in financial savings. Whether the energy savings offset the additional investment costs depends on several aspects, including the extent of the additional costs, the way in which the building is financed (cash or mortgage), interest rates, and the purchase prices of the energy forms over the lifetime of the building. The 1996 building code in Fort Collins, for example, resulted in (estimated) additional investment costs for residences of $1000-1500, while savings amounted to $77-158 per year. Under different circumstances, the period in which the investors would earn their money back ranged from slightly more than 1 year up to 30 years.

[29] See WSU (1997, p. 7), City of Ford Collins (2002, pp. 24-26, 110), WEC (2001, p. 67), Ecofys (2004, p. 39).

2.3.7 Levels of Evaluation Effort

As explained in Chapter 1 (see section 1.3.7), evaluations can be expensive if implemented in detail. Such a comprehensive evaluation is denoted as a 'level A' evaluation. The targeted evaluations (level B) are less rigorous and the programme reviews (level C) are the least intensive forms of evaluation. Table 2.5 gives an overview of the activities that are associated with evaluating regulation measures at each of these levels. Most of the case examples involve a moderate (level B) evaluation. Only the Dutch case, with research into the real energy use, is indicated as level A/B.

According to WSU there are ways to obtain good information on the effects of measures without huge expense, as long as expectations are kept reasonable and the focus is on questions that really matter (thus, questions that really need to be answered). Also, it is not always necessary to know the exact answer; there may be relatively inexpensive pieces of information that are good indicators of success.

2.4 Minimum Equipment Energy Performance Standards Policy Measures and Programmes

2.4.1 Policy Measure Theory

Minimum energy performance standards (MEPS) for end-user equipment are frequently used in combination with labelling. An MEP standard guarantees that the products on the market all perform at least to a minimum level, while the labels stimulate consumers to buy energy-efficient equipment, which should encourage producers to improve the energy efficiency of their equipment. Also, since more and more producers manufacture their models for an international market, this creates the need for international agreements on MEPS.

The theory underlying the MEPS assumes that the performance standard 'cuts off' the worst performing models from the market for a particular product because producers are forced to make models that pass the minimum performance level[30]. This is visualised in Figure 2.6. Before the standard is introduced, the market shows model efficiencies that are typically characterised by the solid-black curve: most models have medium efficiency, only some perform much better, while some perform much worse than average. The performance standard will result in more models that perform above the minimum level (the dashed curve in Figure 2.6). Note that the standard does not motivate producers to make models that perform above average (as far as energy efficiency is concerned). Producers that previously had models with performances below the minimum level will now produce models that are just above the minimum level, instead of producing models that perform significantly better than average. However at the moment the MEPS comes into force governments often announce that the standard will become more stringent in the future. This might cause producers to anticipate this by making models that perform well above the minimum level, which would move the dashed curve in Figure 2.6 more to the right and make it less peaked, as shown by the dotted curve.

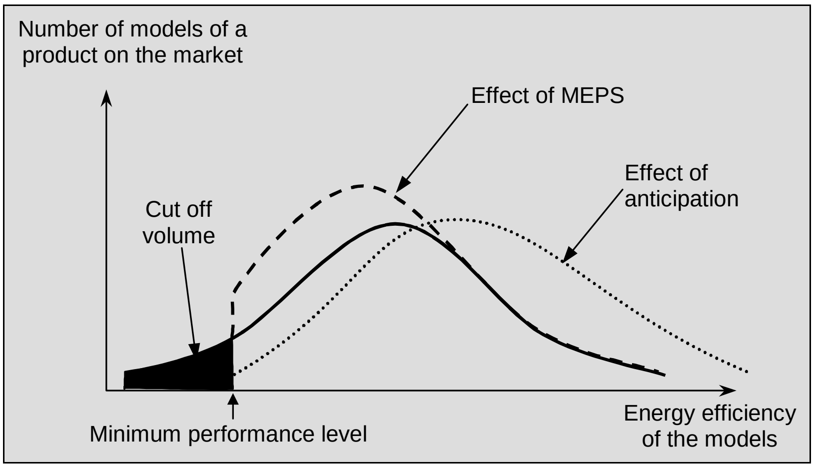
---

[30] See the excellent Guidebook on Energy-Efficiency Labels and Standards by CLASP (2001) and the extensive website on this subject at http://www.clasponline.org/.

Table 2.5 Evaluation activities associated with different levels of effort for building codes

Level A Comprehensive Evaluation	Level B Targeted Evaluation	Level C Programme Review
Quality Control & Analysis of Programme Tracking Data (Inputs And Outputs)		
Literature review Review of programme records, progress reports Expert interviews	Literature review Review of programme records, progress reports	Review of programme records, progress reports
Estimates of Programme Costs		
Review of programme records for administrative costs, costs of modelling heat demand/ energy consumption of buildings, costs of surveys, interviews, information material, etc.	Review of programme records for administrative costs, costs of modelling heat demand/ energy consumption of buildings, costs of surveys, interviews, information material	Review of programme records for administrative costs
Market Characterisation		
Modelling of energy consumption of new buildings Market analysis on number of new-to-build buildings, trends, actors, barriers	Modelling of energy consumption of new buildings Market analysis for baseline estimates	Modelling of energy consumption of new buildings
Development of Baseline		
Ex-ante modelling of energy consumption of new buildings Energy bill analysis Surveys and interviews On-site measuring of energy performance	Ex-ante modelling of energy consumption of new buildings Energy bill analysis Surveys and interviews	Ex-ante modelling of energy consumption of new buildings based on available data
Assessment of Programme Market Effects		
Review of programme records Analysis the enforcement and control of the implementation On-site survey of newly built buildings Market analysis on energy performance of newly built buildings, compliance rate	Review of programme records Quick market analysis based on building applications and/or permits	Review of programme records
Estimate of Gross Energy Savings		
Modelling of gross energy savings Ex-post energy bill analysis Data analysis of on-site measurements, surveys	Modelling of gross energy savings Ex-post energy bill analysis Data analysis of surveys	Ex-post assumptions on energy consumption based on available data Modelling of energy savings
Estimate of Net Energy Savings		
Modelling of net energy savings Data analysis of on-site measurements Data analysis of surveys, interviews including behavioural impacts on energy use	Modelling of net energy savings Data analysis of surveys	
Estimate of Emissions Reduction		
Calculation of emissions reduction with local emissions factors	Calculation of emissions reduction with regional emissions factors	Estimate of emissions reductions with default (inter-national) emissions factors

Figure 2.6 The theory underlying minimum energy performance standards: the models of a particular product that perform below the minimum level are cut off from the market

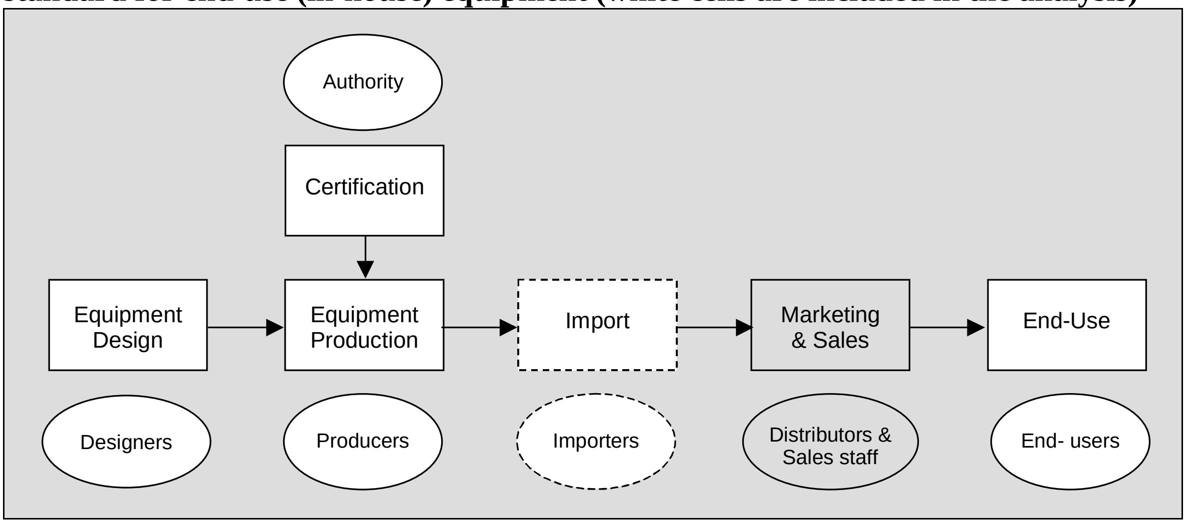


Source: CLASP (2004)

Policy Measure Domain

The life cycle of a model starts at the design stage, followed by the production stage and distribution through marketing and sales. At the end-use stage, the model is actually consuming energy, until it is disposed of. The MEPS will affect the design and production phases, but only for those models that perform below the minimum level. The producers of these models are forced to change their models so that they meet the standard. Nonetheless, all producers will have to use a test protocol to verify that their models comply with the standard. So designers and producers of models are directly affected by the MEPS. Other stakeholders that play a role are, of course, the end-users of the low-efficiency models. Some kind of authority will also have to enforce compliance with the standard. The distributors/ importers and sales staff are not directly affected by the MEPS, but can be affected if the MEPS is accompanied by labelling and other measures (see Chapter 6 on combinations of policy measures). Figure 2.7 gives an overview of the MEPS domain.

Figure 2.7 The stakeholders that are affected by a minimum energy performance standard for end-use (in-house) equipment (white cells are included in the analysis)



Policy Measure Hypotheses

The main goal of minimum energy performance standards is to reduce energy consumption by buildings through the use of more energy-efficient appliances. The governments that implement the MEPS often notify the market that the standard will become increasingly stringent in the future, hoping that designers and producers will anticipate the future situation by making models that perform well above the minimum level. However, awareness of and compliance with the MEPS is essential to achieving energy-efficient appliances.

Other effects of MEPS can include strengthening of competitive markets (the Porter Hypotheses), reduction of the capital investment in the energy supply infrastructure, and enhancement of consumer welfare; but these issues are usually not the driving force behind the introduction of the MEPS, and are not included in our analysis.

Table 2.6 Effects hypotheses on regulation policy measures for various stakeholders in the MEPS domain

Minimum Energy Performance Standards	**Domain specification**	**Hypotheses on outcomes**	**Hypotheses on impacts**
→	**Designers** *Hypotheses on:* Awareness and knowledge of MEPS	→ Energy-efficient designs of appliances	→ Production of energy-efficient appliances
	↑↓	↑↓	
→	**Producers** *Hypotheses on:* Awareness of MEPS Compliance with standards	→ Production of energy-efficient appliances	→ Production of energy-efficient appliances
	↑↓	↑↓	
	Importers *Hypotheses on:* Awareness of MEPS Compliance with standards	Import of appliances that comply with standards	Import of energy-efficient appliances
→	**End-Users** *Hypotheses on:* Awareness of MEPS Awareness of energy efficiency of equipment	→ End-users are triggered to buy (more) energy-efficient appliances	→ Reduced (growth in) electricity consumption compared to baseline

The following section addresses the indicators that can be used to assess the effects of minimum energy performance standards.

2.4.2 *Specification of Indicators*

Indicators that can be used to assess the effects of MEPS include:

- Changes in the level of awareness and knowledge.
- Changes in the adoption levels of practices in designs and models of the product (learning).
- Changes in energy (mainly electricity) consumption of equipment and buildings.

Indicators for the level of awareness and knowledge are addressed in Chapter 3, which focuses on information. Examples of indicators for other categories are listed in Table 2.7.

Table 2.7 Examples of outcome and impact indicators for performance standards

Categories of indicators	**Example indicators**
Level of adoption of practices (learning) in designs and products	• % of energy-efficient designs • energy efficiency of models of a particular product • (changes in) % market share of models
Changes in energy consumption of equipment and buildings	• changed electricity consumption of in-house equipment (kWh/yr) • changed (growth in) electricity consumption for equipment (kWh/yr) • changed (growth in) electricity consumption per building compared to baseline (kWh/yr)

It is important to collect data at the beginning of the process of designing and implementing MEPS, particularly for determining a baseline. A database is often established to monitor the changes in the level of adoption of energy-efficient techniques. This database contains model-specific data concerning annual sales, prices, and technology characteristics of the models. This allows efficiency trends and market shares of models to be monitored. Whenever possible, cooperative agreements with industry should be encouraged, to gather data on sales and efficiency levels. Sales data can be obtained from surveys of manufacturers, retailers, and/or contractors. Note that market share and consumer purchase choices are also influenced by many factors unrelated to relative energy efficiency.

The energy demand of an entire building can be analysed through utility bill analysis, while the energy consumption of in-house equipment can be measured through end-use metering, although the latter is rather costly. Efficiency tests in laboratories with randomly chosen appliances can also be used to check their compliance with the standard.

2.4.3 *Development of Baselines*

The baseline is needed to evaluate the energy savings as a result of the MEPS. To calculate the energy savings, it is important to know not only the energy efficiency of a product or appliance, but also the load at which the appliance is used, as well as the duration of operation. Savings are usually determined by monitoring the energy consumption of a sample of buildings for a full year before the installation of the efficient appliance and then for several years after the installation. However, if loads and operating conditions are constant over time, short-term (e.g. one week) measurements may be sufficient to estimate equipment performance and efficiency. These data are used in engineering models to determine the overall savings. Note that it is important to know whether the new efficient appliance serves as a replacement for older equipment or whether it is used in addition to the older equipment; in the latter case, energy consumption will increase instead of decreasing.

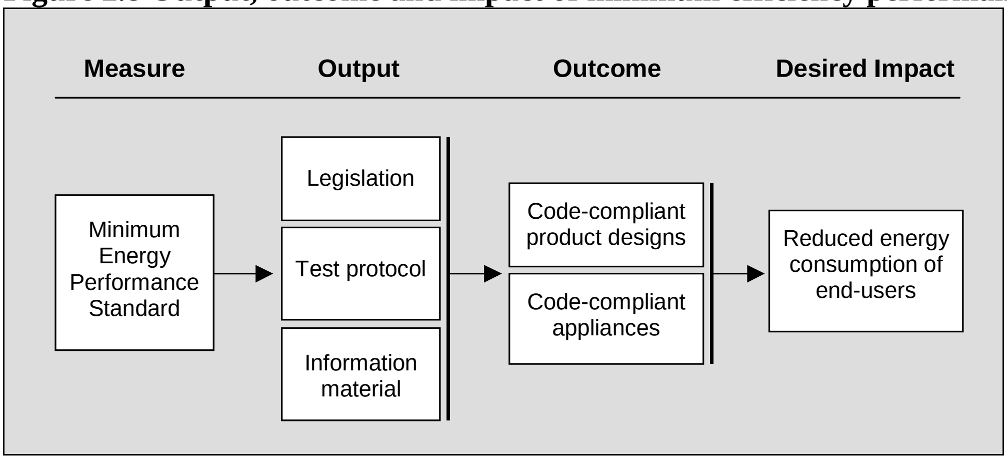
According to the theory discussed in section 2.4.1, the baseline is also represented by the black line in Figure 2.6: if average usage duration of the product is known or estimated, total annual energy consumption of all models of this particular product can also be calculated. LBL (Lawrence Berkeley Laboratory) reports on efficiency improvements in US office equipment (Koomey, 1995) is an example of a baseline model, stock changes and calculated energy savings.

2.4.4 Assessment of Output and Outcome

The output of the MEPS is the legislation in which the standard is described, plus the test protocol that must be used to determine the energy efficiency. The output usually also includes information material to make producers aware of the standard and test protocol (see also Chapter 3 on evaluating information as a policy measure).

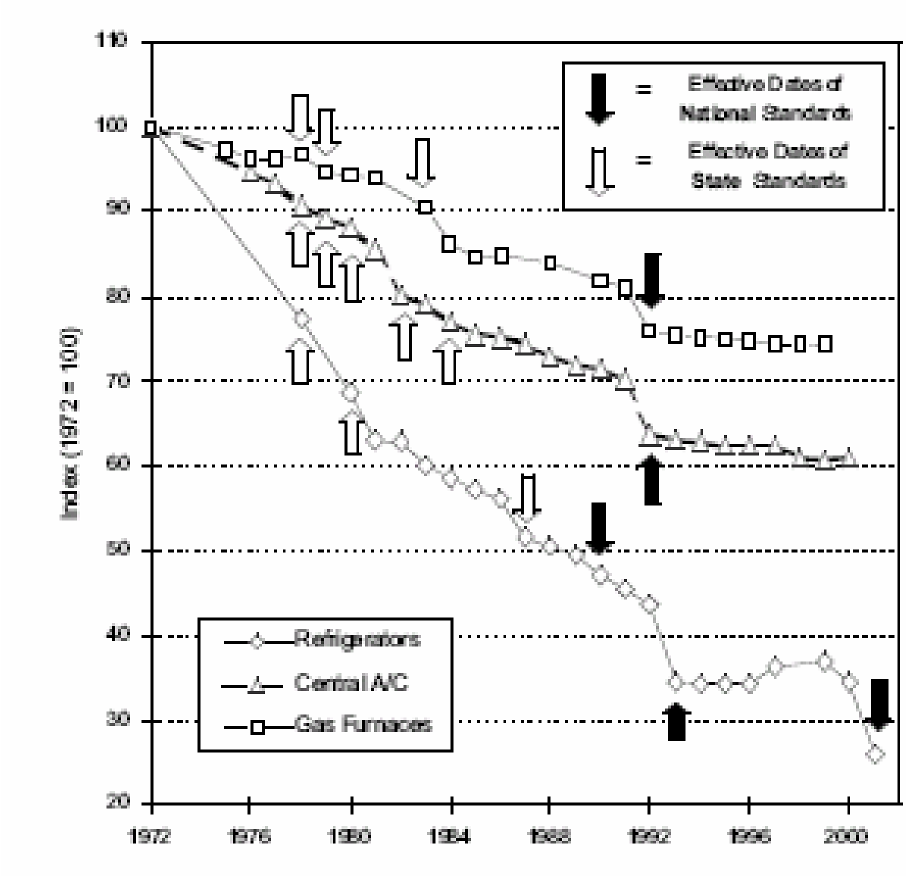
The outcome of the MEPS consists of code-compliant designs that result in code-compliant appliances. Figure 2.8 shows the output and outcome of MEPS that eventually have to lead to the desired reduction in end-use energy consumption. Note that the outcomes of the MEPS, due to the relatively short lifespan of the appliances, become apparent on the market much sooner than, for example, the outcome of building codes. This also implies that evaluations of the performance standards should be implemented before the standards become more stringent, to be able to attribute the effects to the right period. On the other hand, the results of such an evaluation can then be used as input for determining a more stringent minimum performance level.

Figure 2.8 Output, outcome and impact of minimum efficiency performance standards



Another option to present the outcome is the energy intensity of equipment.
Figure 2.9 illustrates that the largest improvements in product efficiency in the US (for three types of appliances) have been in periods adjacent to the effective dates of new efficiency standards.

Figure 2.9 Energy intensity of US refrigerators, central air conditioners and gas furnaces



Source:Nadel, 2003

2.4.5 Assessment of Energy Savings and Emissions Reduction

The energy savings as a result of the MEPS can be deduced from the difference in the curves shown in Figure 2.6, provided that data (e.g. on sales, market share, efficiencies of models) are available to construct the curves *and* that the average load factor and duration of use are known.

The formula is:

$$\Delta E = P_0 \cdot LF_0 - P_t \cdot LF_t$$

Where ΔE = Energy savings [kWh/yr]

P_0, P_t = Nominal power of old (0) or new (t) model [Kw]

LF_0, LF_t = Load factor for old (0) and new (t) model [hrs/yr]: *the (equivalent number of) operating hours in a year that a model operates at nominal power*

Since reduced energy consumption of end-users is the ultimate goal of most MEPS norms, the energy savings can also be deduced from energy bills through proper ex-ante and ex-post analysis. However, both these approaches assume that a new efficient model *replaces* the older model, rather than using the new model as an addition to the older model. In reality, some end-users will continue using their old appliances as long as they still work, and this rebound effect results in an increase in energy consumption instead of a decrease.

On the other hand, if governments notify producers beforehand that standards will become more stringent in the future, this can lead to spillover effects, as some producers will anticipate more stringent standards by making models that perform well above the minimum level.

The Canadian regulations for residential equipment implemented through to the end of 2003 are estimated, by 2020, to have resulted in over 29 megatons per year of GHG emissions reductions. The evaluation of the trends in energy efficiency, as described in Volume II, case example Canada, shows that the average annual energy savings for major appliances is estimated to be 1.56 PJ between 1992 and 2001, with the largest annual saving (of 2.45 PJ) in 2001. On a cumulative basis, between 1992 and 2001, 14.02 PJ less energy was used by major appliances than would otherwise have been the case, had it not been for manufacturers' improvements in energy efficiency and the regulations.

Most appliances need electricity to operate, implying that the emissions reduction due to energy savings can be determined using the shares of the energy sources in the electricity mix and the emissions factors for each energy source (see Chapter 1).

2.4.6 Calculation of Costs, Cost-efficiency and Cost-effectiveness

There is little data available on the costs of implementing the MEPS, but generally the performance standards are believed to be cost-effective instruments for the government (low-cost energy savings) as well as for the consumers that benefit from less energy use. In fact, MEPS are reported to be the single most effective policy for improving the energy efficiency of appliances in the U.S[31]. However, MEPS are not used in isolation: testing procedures and protocols are essential to be able to reliably measure and report the efficiency of appliances on the market. Also important are energy labels, to provide information to consumers and as an incentive to manufacturers to produce more efficient units. In addition, rebate programmes are often designed to encourage and reward producers of models that exceed the minimum efficiency levels by a certain amount. See also Chapter 6 on the combination of policy measures.

The costs of evaluating MEPS (whether or not in combination with labelling) will vary depending on a number of factors[32]. These factors include the quantity and type of available data, and whether energy savings are calculated by engineering estimates or with end-use metering of a sample of products. Most comprehensive evaluations rely on the collection of survey, sales, and billing data. Costs of evaluation will increase if these data are obtained from commercial data collection organisations and/or if end-use monitoring equipment is used to measure energy consumption for specific appliances.

Figure 2.10 presents an overview on the reductions/savings for energy, emissions and costs for three appliances since 1987, and this is an indication for values in this field. For detailed information we refer to Nadel 2003.

---

[31] See DEDE (2002, p. 6); Nadel (2003 p. 80)

[32] See CLASP (2001, p. 155)

Figure 2.10 Estimated savings from US efficiency standards (Nadel 2003)

Enact Year	Standards	Electricity Savings (TWh/yr)			Primary Energy Savings (quads/yr)			Peak Load Reductions (GW)			Carbon Reductions (MMT)			Net Benefit ($billion)
		2000	2010	2020	2000	2010	2020	2000	2010	2020	2000	2010	2020	Thru 2030
1987	NAECA	8.0	40.9	45.2	0.21	0.55	0.61	1.4	14.9	16.5	3.7	10.0	10.1	46.3
1988	Ballasts	18.0	22.8	25.2	0.21	0.27	0.29	5.7	7.1	7.9	4.4	5.0	5.0	8.9
1989&91	NAECA updates	20.0	37.1	41.0	0.23	0.43	0.47	3.6	6.9	7.7	4.8	8.1	8.1	15.2
1992	EPAct (lamps, motors, etc.)	42.0	110.3	121.9	0.59	1.51	1.67	10.1	26.2	28.9	11.8	27.5	27.9	84.2
1997	Refrigerator/freezer update	0.0	13.3	28.0	0.00	0.13	0.28	0.0	1.7	3.6	0.0	2.9	5.5	5.9
1997	Room A/C update	0.0	1.3	2.1	0.00	0.01	0.02	0.0	1.0	1.6	0.0	0.3	0.4	0.6
2000	Ballasts update	0.0	6.2	13.7	0.00	0.06	0.13	0.0	1.8	3.0	0.0	1.3	2.7	2.6
2001	Clothes washer update	0.0	8.0	22.6	0.00	0.11	0.28	0.0	1.3	6.1	0.0	2.2	5.4	15.3
2001	Water heater update	0.0	2.5	4.9	0.00	0.08	0.13	0.0	1.5	3.6	0.0	1.4	2.2	2.0
2001	Central A/C&HP update	0.0	10.7	36.4	0.00	0.11	0.35	0.0	3.5	41.5	0.0	2.3	7.2	5.0
	TOTAL	**88**	**253**	**341**	**1.2**	**3.3**	**4.2**	**21**	**66**	**120**	**25**	**61**	**75**	**186**
	% of projected U.S. use	**2.5%**	**6.5%**	**7.8%**	**1.3%**	**2.9%**	**3.5%**	**2.8%**	**7.6%**	**12.6%**	**1.7%**	**3.4%**	**3.8%**	

2.4.7 Levels of Evaluation Effort

Table 2.8 gives an overview of the activities associated with the three levels of evaluation effort distinguished in this *Guidebook.* The level of ambition in the case examples in Volume II is a moderate (level B) and a programme review (level C).

Table 2.8 Evaluation activities associated with different levels of effort for minimum energy efficiency performance standards

Level A **Comprehensive Evaluation**	**Level B** **Targeted Evaluation**	**Level C** **Programme Review**
Quality Control & Analysis of Programme Tracking Data (Inputs And Outputs)		
Review of programme records, progress reports Collaboration of programme records with stakeholders	Review of programme records, progress reports	Review of programme records, progress reports
Estimates of Programme Costs		
Review of programme records for administrative costs, costs of modelling energy consumption of end-users, costs of surveys, interviews, information material, end-use metering, production costs, etc.	Review of programme records for administrative costs, costs of modelling energy consumption of end-users, costs of surveys, interviews, information material	Review of programme records for administrative costs
Market Characterisation		
Modelling of energy consumption of appliances Market analysis on size, structure, trends, stakeholders, barriers	Modelling of energy consumption of appliances Market analysis for baseline estimates	Modelling of energy consumption of appliances and buildings
Development of Baseline		
Ex-ante energy bill analysis Surveys on energy consumption of end-users On-site measuring of energy performance (end-use metering)	Ex-ante energy bill analysis Surveys on energy consumption of end-users	Ex-ante assumptions on energy consumption based on available data

Level A Comprehensive Evaluation	Level B Targeted Evaluation	Level C Programme Review
Assessment of Programme Market Effects		
Market analysis on size, structure, trends, stakeholders, barriers Surveys with end-users, producers, vendors	Quick market analysis	
Estimate of Gross Energy Savings		
Modelling of gross energy savings Ex-post energy bill analysis Data analysis on-site measurements, surveys	Modelling gross energy savings Ex-post energy bill analysis Data analysis of surveys	Ex-post assumptions on energy consumption based on available data Modelling energy savings
Estimate of Net Energy Savings		
Modelling of net energy savings Data analysis of on-site measurements Data analysis of surveys, interviews	Modelling of net energy savings Data analysis of surveys	
Estimate of Emissions Reduction		
Calculation of emissions reduction with local emissions factors	Calculation of emissions reduction with national emissions factors	Estimate emissions reductions with default emissions factors

As argued in Chapter 1, the level of effort should be selected related to the objective of the evaluation. The above table is a discussion tool on the desired results from the evaluation and to clarify the various options.

2.5 Conclusions

For the evaluation of building codes we concentrate on the codes that relate to heating and cooling demand, and energy performance standards. The domain specification ranges from architects to builders, while the impacts deal with reduced energy consumption. So even if a low level (C) of evaluation effort is selected, there are still options to broaden this or to focus the evaluation, particularly as the programme records will often be related to one specific domain.

Several evaluations (see, for example, Volume II: the Belgium and Dutch case examples) concluded that the enforcement (and control) of the code by local authorities is a weak link in the chain. Staff should have sufficient knowledge, time, and other resources available to carefully review plans and inspect homes. With the increased minimum level the control will become more and more important, as the details of the construction and the correct implementation often mean the difference between compliance and non-compliance. So the importance of the element in an evaluation will increase in the near future, both for the outcome and the assessment of the calculated energy savings.

The impact of a building code becomes apparent only after a considerable period of time due to the time lag that exists between permit application and actual construction of a building. In addition, recurrent revisions of a building code make it difficult to pinpoint effects within a certain period. The increased demand for comfort results in a trend towards larger buildings, and thus larger spaces to be heated or cooled, leading in turn to increased energy consumption. The rebound effect also results in increased energy consumption as users change their behaviour when faced with new technologies.

MEPS are almost always used in combination with other policy measures such as labelling and rebate programmes to improve the desired impact of the measure (see also Chapter 6). If MEPS are used on their own, only the worst performing models are cut off from the market, as the mean of the model-performance curve will not change. A Level-A evaluation is therefore hard to justify.

The energy savings that could be achieved with more efficient appliances are often offset by the fact that end-users do not replace their old appliances, but use the new efficient appliances alongside the older ones, thereby causing an increase in energy consumption instead of a decrease. This rebound effect (replacement versus additional use) should be incorporated as a clear assumption in the models that are used to estimate and assess the energy savings.

3. EVALUATION OF INFORMATION POLICY MEASURES AND PROGRAMMES

3.1 Introduction

The majority of the case examples included in Volume II are evaluations dealing with information dissemination. Table 3.1 presents the cases organised in the subcategories used for information policy measures.

Table 3.1 Evaluation case examples dealing with information included in Volume II

Subcategories	Case examples	Country
General information programmes	The ‘A’ campaign 1999	Denmark
	Promotional campaign for efficient ventilation	Denmark
	Information campaign (2001)	France
	Energuide for houses	Canada
Labelling	Energy Efficiency Rating Labelling	Korea
	Energy Efficiency Rating Labelling	Korea
Information centres	Local energy efficiency information centre (Espaces Info Energie, EIE)	France
	Information centres in local region	Sweden
	Local energy efficiency information centres	Belgium
Energy audits	Energy labelling of small buildings	Denmark
	Free-of-charge electricity audit	Denmark
	Audits (“Aides a la decision”)	France
	Energy Audits in industry	Korea
	Energy Audits in buildings	Korea
Education and training	Project ‘Red-Hot’ (element of stand-by campaign)	Denmark
	Information and educational programme 1998-2002	Sweden
Demonstration		
Governing by example		

Information from these case examples is used in the relevant sections to illustrate elements such as output indicators, development of baselines and energy savings.

This chapter begins with the objectives and principal kinds of information policy measures and programmes. For the subcategories (reorganised into three groups) we present the main evaluation topics, structured according to the key analytic elements as presented in Chapter 1:

Element 1: Policy measure theory used
Element 2: Specification of indicators for the success of a measure
Element 3: The baselines for the selected indicators
Element 4: Assessment of outputs and outcomes
Element 5: Assessment of energy savings and emissions reductions and other relevant impacts
Element 6: The calculation of costs, cost-efficiency and cost-effectiveness
Element 7: The level of evaluation effort

Based on the country examples and the discussions at the experts meeting, it is clear that the following issues are of considerable interest for evaluating regulation policy measures and programmes:

- At the theory level (and particularly the hypotheses) specific attention should be given to the barriers: to what level will these be removed?
- One should look carefully into the persistence of the information. This is particularly important for the relationship between output and impact (indicators).
- In most situations the impacts are hard (specific) or almost immeasurable, with a distinctive relation to the policy measures. However, this is not the case for cost efficiency.
- The timing of the evaluation differs for the various subcategories. For example, the outcome of advertising can be measured immediately after finalisation, but the outcome of education only becomes apparent after a period of time.

3.2 Objectives and Main Types of Information Policy Measures and Programmes

In general the objectives of information policy measures and programmes are to:

- Increase the awareness of consumers, manufacturers, vendors, and 'intermediate market stakeholders' such as architects, engineers, and equipment distributors of energy-efficient products and services, as well as their economic and environmental benefits.
- Persuade consumers and vendors to adopt energy-efficient products and practices.
- Provide vendors and, in some cases, consumers with the technical information they need to identify and adopt energy-efficient practices.

The evaluation literature identifies the following types of information-based energy efficiency programmes.

General Information. These policy measures and programmes consist of paid advertising (in newspapers, television, radio, flyers etc.) and public-relations campaigns (as energy weeks, conferences and mailings). In most cases they are designed to make consumers aware of the need to save energy, the means at their disposal for doing so, and the consequences of not doing so.

Labelling. Labelling policies, as implemented in the European Energy Label Programme, generally consist of a number of coordinated components. The first is a series of structured negotiations with equipment manufacturers on developing a labelling system. An important phase is the development of energy efficiency standards: these are a set of procedures and regulations that prescribe the energy performance of manufactured products. The next step is the specification for qualifying equipment, grades or levels of qualifying equipment, efficiency testing procedures, label design, and rules for the placement of labels. The second component consists of advertising, merchandising, and public-relations efforts to encourage vendors to promote qualifying products and to encourage consumers to buy them.

Information Centres. Information centres are developed specifically to package and disseminate information on energy-efficient products and practices. The information generally consists of fairly technical materials designed to orient vendors and consumers towards energy efficiency opportunities, as well as materials such as case studies designed to encourage market stakeholders to pursue those opportunities. Often, the work of these centres is oriented towards a defined set of technologies, markets, or groups of market stakeholders.

Energy Audits. Energy audits consist of a structured inspection of a facility to estimate energy use and identify opportunities for increasing energy efficiency. In some cases, the customer himself carries out the inspection using protocols developed by the programme manager. On-site observations are analysed to allocate metered facility energy use to specific end-uses, estimate savings associated with applicable efficiency measures, estimate the costs of those measures, and prepare investment analyses of those measures. Energy audits are designed to help facility owners overcome a number of common barriers to implementing energy efficiency measures. These include the reduction of information costs, mitigation of information asymmetries (by providing economic analysis of potential measures from a party with no financial interest in their implementation), and reduction in perceptions of risk.

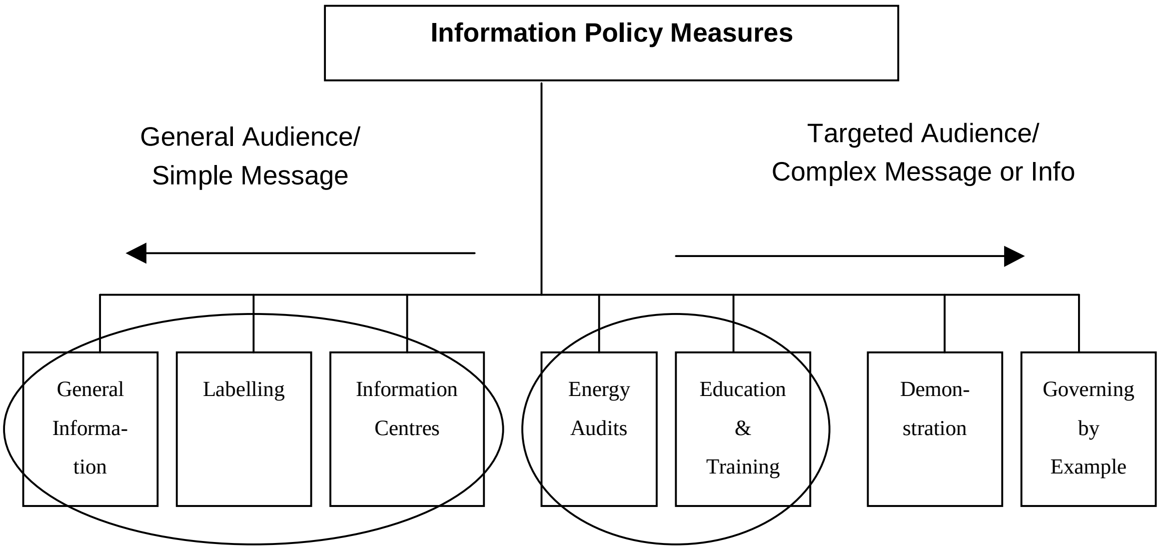
Education and training. Education and training provide focused information on energy efficiency opportunities, particularly end-use technologies or clusters of technologies. These programmes are often targeted towards market parties in the supply chain: equipment vendors, engineers, and design professionals. In some cases, training programmes are associated with professional or trade certification, but also with future energy users (e.g. schoolchildren).

Demonstration refers to the phase during which this new product or technique is tested in practice. This serves to generate information on the usefulness, costs and energy savings during real use or to demonstrate this product or technique to potential users or decision-makers. Pilot programmes are not included in this category, but in the group of policy measures that they are demonstrating.

Governing by Example. A demonstration of practical use is often used to highlight information on a new technology. Target groups, at whom this new technology is aimed, are invited to visit the location where the implementation is being demonstrated. Once on-site they are given additional information and documentation, etc. Sometimes governments (e.g. Belgium and the Netherlands) choose their own governmental buildings, appliance purchasing schemes, etc. to demonstrate energy savings.

Figure 3.1 illustrates the range of information dissemination, from a simple message up to more complex information, combined with information for a general audience through to specific target group.

Figure 3.1 Varieties of information dissemination programmes



To facilitate the following exposition, the policy subcategories have been organised into three groups. The first consists of general information, labelling and information centres. These types of programmes are broadly targeted towards the general consumer through mass media and public relations, and generally convey either a very broad message ('saving energy is good for the environment') or endorsement of simple actions ('buy a labelled appliance'). Information centres are often targeted towards broader markets, but also sometimes towards a specific group (such as auditors), or towards a specific technology (such as heat pumps etc.). Energy Audits and Education & Training, by contrast, attempt to convey more complex technical information to more specific audiences. Typical examples include resource centres on energy efficiency for architects, or training in using variable speed drives in ventilation systems. Demonstration projects for energy efficiency measures, concepts, buying groups etc. by local, regional or national governments are examples of a more complex message for a targeted audience.

3.3 General Information, Labelling and Information Centres

3.3.1 Policy Measure Theory

The policy measure theory provides the basic framework for the evaluation as it identifies the relevant market parties (the policy measure domain) and the hypotheses effects.

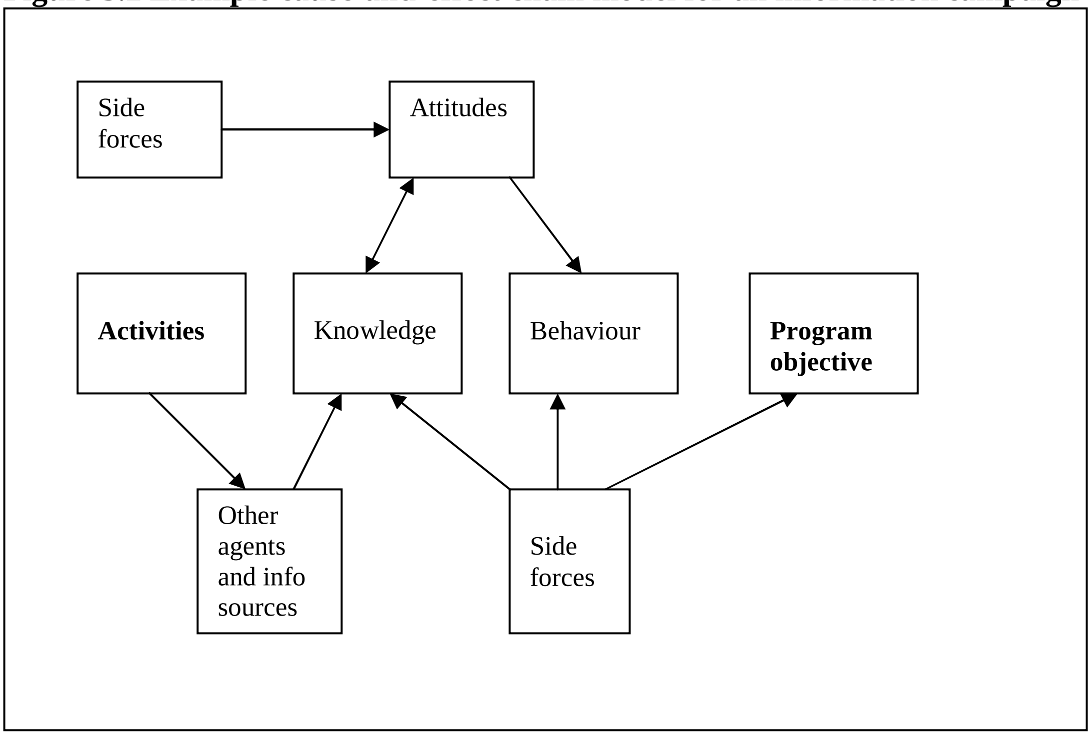
Policy Measure Domain. Because general information and labelling programmes rely primarily on mass-media advertising and public-relations activities to achieve their objectives, their appropriate domain is the consumer population. Generally, the distinction can be made between domestic and business consumers.

Policy Measure Hypotheses. The basic hypothesis concerning the effects of general information, labelling and information centres programmes is as follows:

1. Consumers are inhibited from adopting cost-effective, readily available efficient products and practices due to lack of information; this might be specified in a number of barriers relating to awareness, understanding or knowledge.
2. The missing information may include one or more of the following: energy savings, other performance benefits, pricing, related environmental burden, product availability (compared to conventional technologies or practices).
3. Obtaining this information entails costs: expenditure of time and effort.
4. Other conditions may impose additional information costs. These include behaviours and attitudes developed under earlier technical circumstances and conflicting information obtained from market parties with different interests in the subject technologies.
5. Effective information campaigns reduce consumer costs and perceived risks of adopting efficient products and practices, and thereby increase the adoption pace beyond what would have occurred in the absence of the programme.

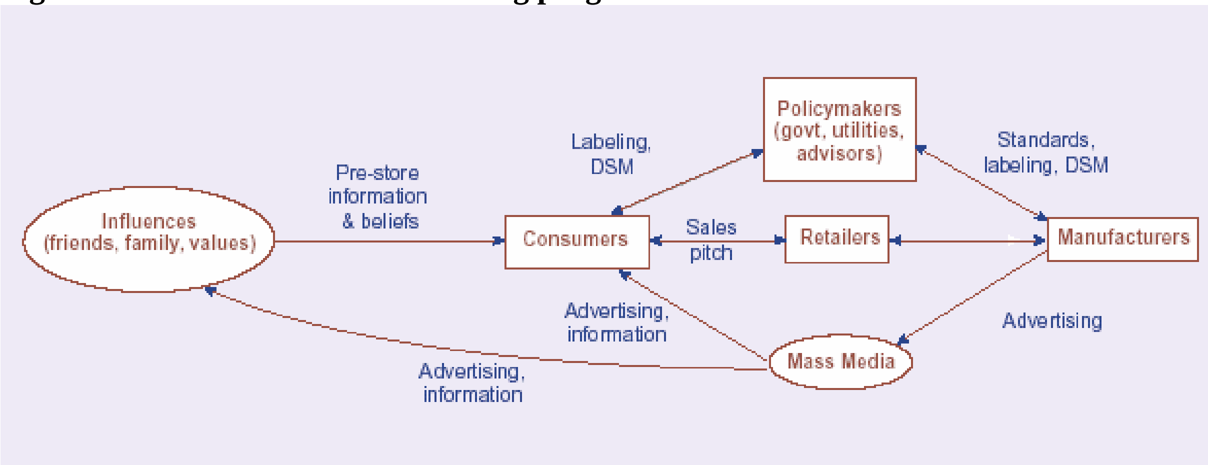
Most hypotheses dealing with general information use some kind of general causal model, including knowledge, attitudes and behaviour as main features. Figure 3.2 shows an example used in Scandinavia. Figure 3.3 shows a hypothesis from another point of view: a stakeholder model for labelling.

Figure 3.2 Example cause-and-effect chain model for an information campaign



Source: Motiva, 1999 page 62

Figure3.3 Stakeholders in the labelling programme



Source: Wier, 2001

3.3.2 Specification of Indicators

Input indicators refer to the money used for creating and distributing information (general information, label information or products in information centres) and man-hour capacity.

Output indicators refer to the countable products such as the number of television commercials, number and regions of distributed leaflets, hits on a website, visits to information centres, etc.

Outcome. As Figure 3.2 suggests, indicators of programme outcome for general information, labelling and information centre programmes will focus on:

- Changes in customer awareness levels and attitudes towards energy efficiency.
- Changes in knowledge levels of efficient products and practices.
- Better understanding of the meaning of efficiency labels and ratings.
- Use of labels and ratings in purchasing decisions.
- Adoption (change of behaviour) of the targeted practices and products.

The policy measures can also include targets other than energy savings and, in those cases, additional outcome indicators could be selected. This *Guidebook* is limited to indicators relevant to energy.

Figure 3.3 shows how the various stakeholders interact and affect the purchasing environment, and ultimately, the consumer's purchasing decision. Evaluators initially focus on outcome indicators – changes in attitudes and behaviour of market players ('leading indicators'), which can be measured in shorter periods of time – rather than energy savings, appliance sales, and GHG emissions reductions ('lagging indicators')[33]. Table 3.2 contains examples of outcome indicators for labelling programmes.

[33] Wiel, S. and McMahon, J, Energy-Efficiency Labels and Standards, 2001, pages 153-155

Table 3.2. Example of outcome indicators for labelling programmes

Relating to consumers

- Level of awareness of the energy label, related product material and advertising
- Degree of influence that the label has on purchase decision; how does the consumer understand the label?
- Increased sales of efficient equipment
- More frequent recommendations or specification of energy-efficient equipment and design

Relating to retailers

- Sales
- Attendance at training courses and intent to implement training
- Changes in costs of efficient equipment
- Changes in equipment stocked or displayed by retailers
- Attitudes of retailers
- Increases in private sector advertising in support of efficient technology

Relating to manufacturers

- Sales
- Increased knowledge or awareness among planners, designers and decision makers about energy-efficient technologies
- Changes in costs of efficient equipment
- Direct and indirect costs to manufacturers (cost of production, R&D efforts to improve appliance efficiency, distribution of labels, promotion and support of labelling programmes)

Adapted from Wiel, 2001; Banks, 2002

Volume II contains case examples. Table 3.3 contains a scheme from the Danish A-campaign that illustrates the relationship between the (expected) output of the labelling campaign, selected outcome indicators and the energy and emissions impacts.

Table 3.3 Danish labelling campaign in1999 and selected indicators

Produced output	Some assessed outcome indicators	Energy impact
➢ In-store and trade material ➢ Rebate system ➢ Public announcements (TV, radio, press releases, other advertising) ➢ Web-list of A-label appliances	➢ No. of customers who were motivated to buy by the rebate ➢ No. of customers who know the website ➢ No. of customers who feel that the recommendation by the electricity company was important to their purchase ➢ Sales of A-label appliances, by type ➢ Market shares of A-label appliances, by type ➢ Product range, by type ➢ Price reductions, by type ➢ Hits on website ➢ No. of retailers who report to web-list ➢ Supply compared to demand ➢ Retailer satisfaction	➢ Realised electricity savings ➢ CO_2 reduction ➢ (CO_2 abatement cost)

Evaluations of information programmes have identified a number of other programme design features and market conditions that affect the outcomes of information programmes. These include:

- Level of recognition and credibility of the sponsoring agencies.
- Number and diversity of channels and sources of programme information.

- Rate of message dissemination (number of impressions, frequency of press runs).
- Clarity of the message.
- Adoptability of the message.

The **impact** indicator is always changing in energy use and related emissions.
For labelling policy measures the outcome indicator related to sales is often (wrongly) presented as an impact indicator. The reason for this is the fact that market share information is often not immediately available during the implementation phase.

3.3.3 Development of Baselines

General information policy measures include assumptions that at some point the changes in awareness level, knowledge and adoption level will result in energy savings. There are models to calculate expected energy savings, but the assumptions regarding the relationship between the changes in levels and savings have the dominant impact.

For general information and energy centres, the main elements for a baseline study are:

- Size and composition of target markets or audience.
- Pre-programme awareness of knowledge levels.
- Pre-programme information and education sources.
- Extent of exposure to (and use of) pre-programme education or information sources.
- Pre-programme status of the target market relative to the intended results of the programmes.
- Pre-programme adoption patterns.

Benchmarking could form an alternative (or an enrichment) to the baseline for general information campaigns. The outcomes of other national public (advertising) campaigns (e.g. on healthcare or new cars) are then used as a reference and for reference values.

In the case of labelling policy measures, stock models are used to determine the baseline. These stock models use retail sales and appliance attribute data to estimate the average efficiency of new appliances sold in any one year and then tracks appliances through to their eventual retirement and scrapping by making assumptions about average appliance service life. In order to measure the impact, a baseline estimate is constructed for a business-as-usual (BAU) scenario, against which the actual and projected efficiency trends can be compared. A popular method for establishing the baseline is to use a number of independent estimates and then make an informed judgement on the baseline.

Information Centres. The nature of most information centre programmes is to have relatively low-intensity contact with a large number of users. Therefore telephone, mail, and Internet surveys have proven the most effective means of collecting information concerning changes in programme effects indicators among programme participants. From a conceptual standpoint, it would be best to capture information on participants' baseline levels of efficient product knowledge and adoption at the time they enrol in or first use the programme. However, there are no instances in the literature in which such an approach has been taken. The author's personal experience is that programme managers resist efforts to collect 'intake' data for fear of discouraging use of the programme and increasing administrative costs.

3.3.4 Assessment of Output and Outcome

Table 3.4 shows examples of effects indicators that have been used for general information and labelling programmes.

Table 3..4 Examples of outcome indicators for general information, labelling and information centres policy measures

Outcome	Examples
Change in awareness level	• Δ (over time) percentage of consumers that can name efficient practices (in regard to a specific end-use). • Δ percentage of consumers who have seen efficiency label prior to a survey, or who can correctly describe the label without visual prompt.
Change in knowledge level	• Δ percentage of consumers who can accurately describe energy and performance advantages of efficient products. • Δ percentage of consumers who can accurately describe cost differences between efficient and conventional products. • Δ percentage of consumers who report that energy efficiency has a high priority among product features.
Change in adoption level of practices	• Δ percentage of consumers who report efficient practices. • Δ percentage of consumers with practices observed by vendors or independent on-site inspection.
Change in adoption level of efficient products	• Δ in market share of efficient models or technologies for targeted end-uses.

Verification of Output and Outcomes. Verification of outcomes for general information and labelling programmes poses unusual challenges for energy efficiency programme evaluators. Programme activities consist of broadcast messages or, in the case of label placement and point-of-purchase display, include diversified activities that cannot be managed centrally. Programme managers and evaluators have used the following data collection strategies to verify the output impacts of general information and labelling programmes.

- ***Media market statistics.*** Managers and evaluators of the US Environmental Protection Agency's ENERGY STAR® Program have relied extensively on media market statistics to quantify the outcome of the programme. These consist of independently audited circulation, listener, and viewer counts for paid advertising. The agency also pays media monitoring services to collect mentions of the programme resulting from press releases and other public-relations activities.

These types of statistics were also used to evaluate the French national information campaign (included in Volume II), and it was concluded that 86% of the target population was reached at least once, with an average rate of 3 'exposures' to the message. The total number of contacts was estimated at 157 million. In addition, 84% of the respondents considered the message important and 76% viewed it as credible.

- ***Retail location inventories.*** Much of the early evaluation efforts for the European Energy Label focused on assessing the extent to which consumers encountered the label when shopping for appliances, as fully specified in the relevant EU regulations. Data collection for this analysis consisted mainly of primary observation of a representative random sample of retailer locations and websites for distance selling. Indicators developed from these observations included the percentage of displayed appliances (of different kinds) that carried fully compliant label information, partially compliant, and none at all. An early evaluation found that the percentage of appliances with fully compliant labelling ranged from 7-95% in the 15 EU nations.[34]

Evaluation is often restricted to output indicators. For example the evaluation of Swedish Municipal Information Centres on energy issues has been limited to the use of output indicators such as: number of municipal energy information centres, activities at the information centre (telephone information, seminars etc.), type of information requested, the use of the websites at the municipal energy information centres etc. Volume II contains more information on this case example.

Samples to assess changes in indicators. Given the nature of general information and labelling programmes, developing information on outcome indicators is usually accomplished through large-scale population-based, random-sample surveys. Sample sizes of 1,000 to 1,500 have been used for surveys covering national populations. Samples of this size support a high level of segmentation and relatively precise estimates of sample proportions and means. Evaluators of such programmes have used mail, telephone, and Internet survey administration techniques to gather information on product and programme awareness, product knowledge, reported adoption of efficient products and practices, and programme influence. For Type B (Targeted Evaluations) and Type C (Programme Reviews), surveys with significantly smaller sample sizes will generally be adequate.

Evaluators have also taken a wide range of approaches to developing data on the market share of efficient products and prevalence of efficient practices.

Methods to assess outcomes. Evaluators have generally relied on the following approaches to assess the outcome of general information and labelling programmes.

- ***Historical analysis:*** implementation of broad-based consumer surveys at various stages of programme development. Best practice in this regard is to keep the survey text and sampling methods the same (or as similar as circumstances will allow) from one phase to the next. This approach best captures the change in key indicators over time.
- ***Cross-sectional analysis.*** In some cases, information programmes are implemented differently from one region to another or are rolled out in phases. This provides the opportunity to implement cross-sectional analysis through the proper structuring of samples and hypotheses.[35] In other cases, information on developing key variables is available from comparable areas where there are no programmes (or very different programmes).

---

[34] Windward, John et al., (1998). *Cool Labels: The first three years of the European Energy Label.* Oxford: Environmental Change Unit, University of Oxford.

[35] See, for example: Goldberg, Miriam L. et al., 'Counting the Stars in America's Eyes: The Energy Star Household Survey', in *Proceedings of the 2001 International Energy Program Evaluation Conference.* Salt Lake City, 2001.

- ***Self-reports.*** Many evaluations of information centre programmes make extensive use of participant surveys to develop programme effect indicators. Because it is difficult in practice to obtain baseline information as part of the programme enrolment process, many evaluations have taken the approach of questioning participants about practices prior to and after participation in the programme (in the same questionnaire). One way to mitigate the self-reporting bias that is related to this method of research is to use open-ended questions to probe whether customers have changed opinions after receiving information from the centre.

The French case example (in Volume II) presents results from the use of the contact database that each information centre maintains. Analysis showed that, since 2001, the local centres have received 80,000 contacts, of which 84% came from households, 9% from organisations, and 4% from building professionals. Around 60% are telephone contacts, 30% from a visit to the centre and 10% via e-mail. The most frequent questions concern projects relating to residential improvements (41%), then new homes (29%) and the retrofitting of residences (23%). The demands addressed to the info centres by the different types of consumers or stakeholders vary considerably. For households, questions generally concern space heating and domestic hot water (both around 60%), then energy management (25%), insulation (20%), lighting (10%) and electrical appliances (7%).

3.3.5 Assessment of Energy Savings and Emissions Reductions

Analysis of Energy Savings. Where analysts have estimated energy savings attributable to general information programmes or the consumer education elements of labelling programmes, they have done so by using models. For example, by first estimating the number of efficient units sold (or the number of consumers adopting efficient practices) as a result of the programme, then multiplying the 'net adoptions' by an estimate of unit energy savings. This latter estimate is generally developed through engineering analysis or building modelling.

Readers should bear in mind that the design of some information programmes might preclude or complicate estimates of net adoptions and/or unit energy savings. Examples of such situations include instances where the programme:

- Broadly promotes a label or concept without calling for specific actions or purchase of specific items.
- Promotes actions such as turning off unnecessary lights, where it is difficult to estimate unit energy savings.
- Works in coordination with fiscal policy measures (rebates), which provide a much stronger incentive for consumers to adopt efficient products and practices.

Some authors indicate that labelling programmes require at least a 1-2 year period for the follow-up and give time for policy changes to spread to all market parties and for changes to become embedded.

In most cases, especially for general information campaigns and for information centres and education and training schemes, it is difficult (or almost impossible) to make a good estimation of the energy efficiency improvements (impacts). For the labelling policy measures there are two options: to use the energy savings data from the test procedures or to measure the energy use during actual use.

3.3.6 Calculation of Costs, Cost-efficiency and Cost-effectiveness

Costs relating to general information programmes and labelling programmes normally concentrate on cost efficiency, i.e. the ratio between the input and outputs.

3.3.7 Levels of Evaluation Effort

As discussed earlier, general information and labelling programmes are characterised by a large volume of highly diversified activities and consumer contacts (impressions). Therefore, evaluation efforts must begin with systematic efforts to record and, wherever possible, quantify these outreach and publicity activities. This is very difficult to do retrospectively. Comprehensive evaluations will require a baseline consumer survey with a relatively large, population-based random sample, covering awareness and adoption of the targeted behaviours or products. These should be followed up at annual intervals to track progress. Finally, another baselining strategy, such as cross-sectional analysis or consumer self-reports, should be developed and used in order to assess net effects of the programme. Targeted evaluations will generally consist of fewer surveys or may substitute lower cost data collection mechanisms, such as focus groups for full-scale sample surveys. See Table 3.5.

Table 3.5 Evaluation activities associated with different levels of effort: general information and labelling programmes

Level A **Comprehensive Evaluation**	**Level B** **Targeted Evaluation**	**Level C** **Programme Review**
Development of programme theory and specification of indicators		
Literature review. Expert interviews. Consumer focus groups.	Literature review. Expert interviews.	Literature review.
Characterisation of programme activity		
Quality control of programme records, particularly advertising reports. Analysis of programme records. Corroboration of records with programme staff and contractors.	Analysis of programme records. Corroboration of records with programme staff and contractors.	Analysis of programme records.
Estimation of changes in indicators		
Population-based sample surveys: prior to programme launch, then annually.	Population-based sample surveys – lower frequency or smaller samples. Focus groups	Focus groups. Interviews with retailers, vendors Expert interviews.
Baseline development/estimation of net impacts		
Cross-sectional analysis. Survey of consumers in areas without programmes or with different types of	*Cross-sectional:* Use of primary survey results from other areas. *Self-reports.* Consumer surveys covering self-reports of	Focus groups. Interviews with retailers, vendors. Expert interviews.

Level A Comprehensive Evaluation	Level B Targeted Evaluation	Level C Programme Review
programmes. *Self-reports.* Consumer surveys covering self-reports of programme effects. *Historical Time-Series: Analysis* of changes in efficient product/practice market share.	programme effects. *Historical Time-Series: Analysis* of changes in efficient product/practice market share.	
Estimates of energy savings		
Engineering analysis. Engineering analysis with primary data collection on parameter values.	Engineering analysis.	Engineering analysis.

The case example in Volume II, on evaluations for information centres, general information campaigns and labelling policy measures, includes the whole range of effort (A, B and C levels) and shows that it is important to choose the appropriate level.

3.4 Energy Audits and Education & Training

3.4.1 Policy Measure Theory

Policy Measure Domain. Energy audits and training programmes are generally targeted towards specific groups of customers defined by industrial branch or the presence of specific end-uses in their facilities. The domain for these programmes also includes market parties in the supply chain of equipment or installations typically used by the targeted customers: manufacturers, engineers, architects, distributors, installers, and maintenance contractors. Evaluations of training programmes have faced a particular challenge in defining and identifying the businesses on the supply side of the domain. The problems arise from a number of sources including: geographic dispersion of firms providing relevant services to energy consumers in the programme area, plus unclear or overlapping membership in trades, professions, and their representative associations.

Policy Measure Hypotheses. The basic hypotheses are as follows.

1. Facility owners face a number of *information-related barriers* to adopting energy-efficient technologies. First, facility owners and their managers are not generally aware of the broad range of energy efficiency opportunities available to them. Second, information on energy efficiency measures for complex building and production systems is generally given to facility owners by equipment vendors or engineers with a financial interest in selling a particular technical solution. Thus, it is difficult and costly to obtain an ‘objective’ view of the costs and benefits of a given technical solution for the facility in question. Third, facility owners and their management staff frequently lack the broader technical background to place the claims of competing technologies in perspective or to compare the relative merits of those technologies. Fourth, facility owners and managers perceive that there is a certain risk in adopting new technologies or approaches that may compromise the performance of critical building or production systems. Finally, information that could help facility managers to address the barriers

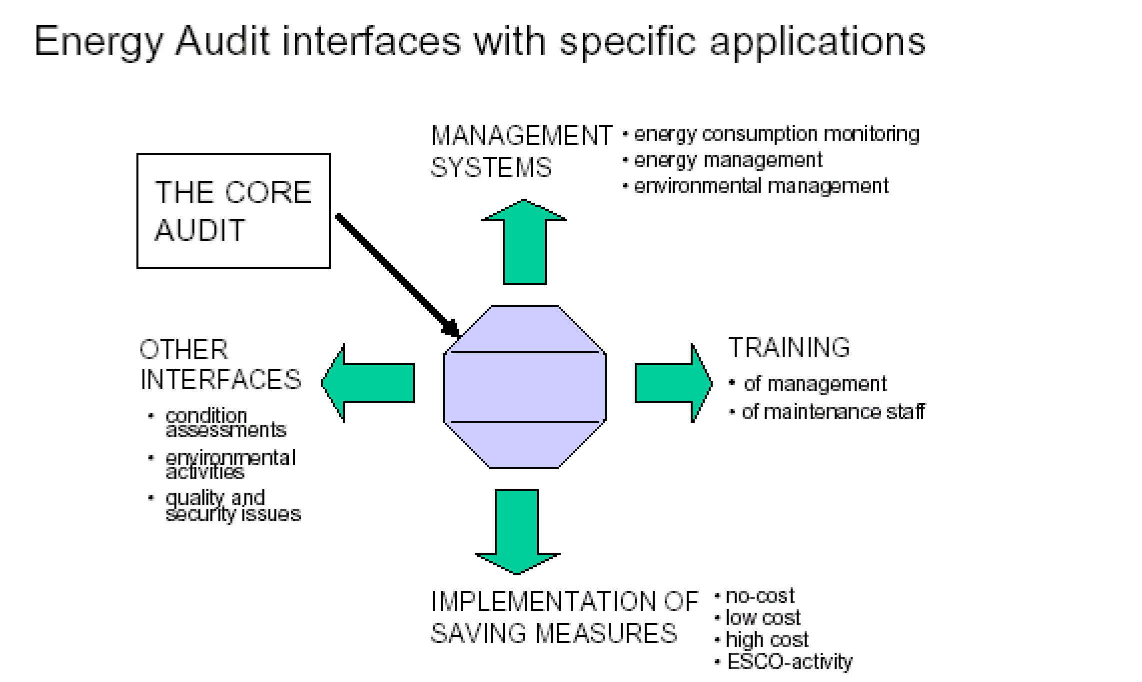
identified above tends to be scattered between many, often obscure, sources. Thus, the costs of overcoming the barriers are high.

2. Engineers, designers, vendors, and installers in the supply chain for building and production systems face a related set of *barriers to consistent promotion and use of energy-efficient products and design approaches*. These include costs associated with obtaining the necessary technical information on technology performance and applications, 'objective' assessment of competing products and approaches, and perception of performance risks. Moreover, many market parties in the supply chain operate in highly cost-competitive sectors. One consequence of this economic environment is a reluctance to take on the 'learning costs' associated with using new technologies.

3. The three categories of policy measures address these barriers through the following activities.
   - Compilation and indexing of technical information from various sources for easy access.
   - Development of case studies and other materials to mitigate perceptions of risk. In the case of energy audits this is the central item.
   - Focused training sessions for facility users, vendors, designers, and engineers.
   - Operation of 'tool lending libraries' for measurement equipment and computerised measurement assessment programmes.
   - Information dissemination through various channels: direct mail, e-mail, print, lectures.

An example hypothesis of interrelations between energy audits and specific applications is shown in Figure 3.4

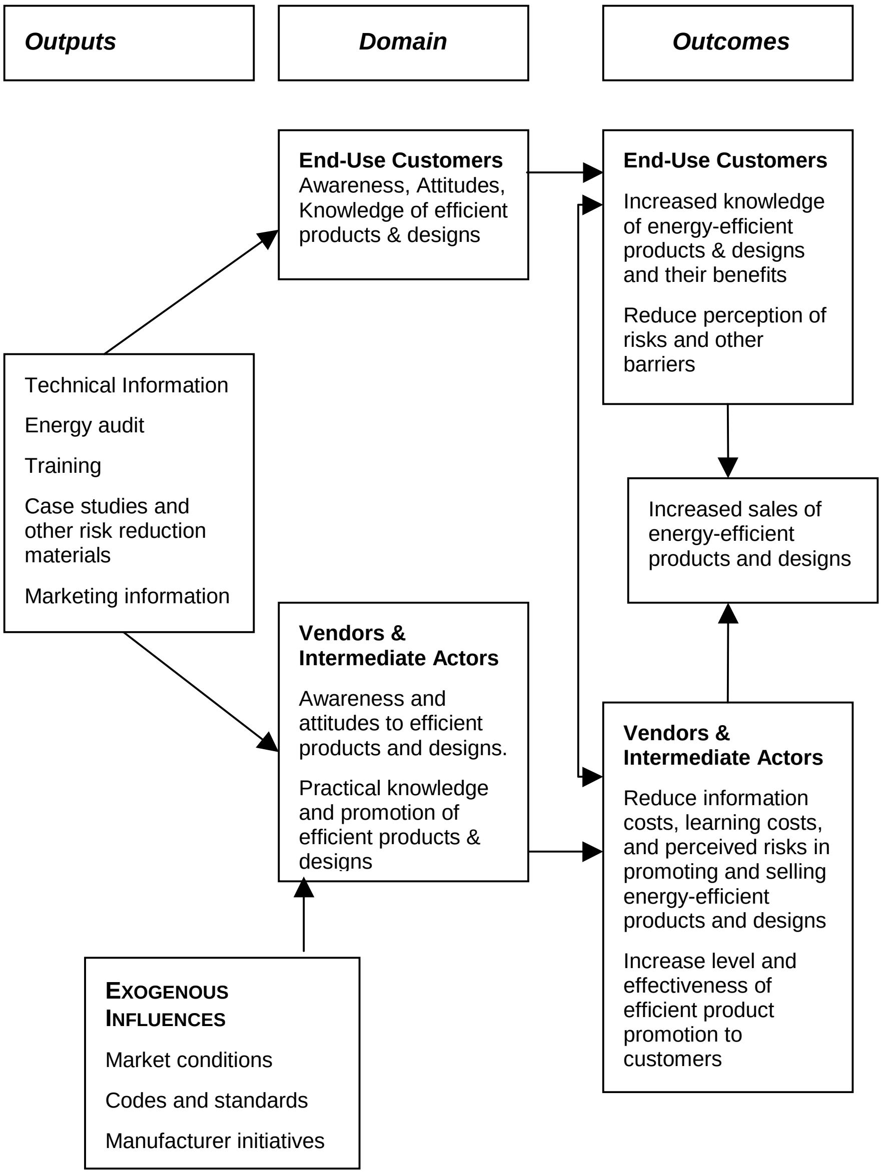
Figure 3..4 Example hypothesis of interrelations between energy audits and specific applications



Source: Presentation of SAVE project, Energy Audit management procedures, page 12

Figure 3.5 shows an example of a general causal model for these three categories of policy measures.

Figure 3..5 Model of education and training effects



3.4.2 Specification of Indicators

As Figure 3.5 suggests, indicators will focus on **outcomes**, i.e. changes in customer awareness levels, knowledge, and adoption of efficient products and practices. For vendors, designers, and other intermediate parties, outcomes will primarily consist of increased knowledge of energy-efficient products and designs, increased confidence in their performance, and increased levels of effort and effectiveness in promoting those items to customers. For both end-users and market parties in the supply chain, recognition and increased use of the information centres, energy audits or training programmes will also be important indicators of programme effects. Table 3.6 shows examples of outcome indicators that have been used in evaluations.

Table 3.6 Examples of outcome indicators for energy audits and education & training policy measures

Outcomes	**Examples**
End-Use Customers	
Change in awareness level	• Percentage of targeted customers who are aware of the programme. • Percentage of advised measures in the audit discussed by the management • Δ percentage of consumers who can name efficient products and practices (in regard to the targeted end-uses).
Change in knowledge level	• Δ percentage of consumers who can accurately describe energy and performance advantages of efficient products and design practices. • Δ percentage of consumers who report that energy efficiency has a high priority for targeted products and system designs. • Percentage of advised measures with an acceptable payback period
Change in adoption level of energy-efficient products and practices	• Δ percentage of consumers who report efficient practices. • Δ percentage of consumers with practices observed by vendors or independent on-site inspection. • Δ percentage of consumers who report adoption of purchasing policies that favour or require efficient products or designs. • Percentage of advised measures implemented within three years after the recommendation.
Market Parties in Supply Chain	
Change in awareness level	• Percentage of targeted supply-side parties that are aware of the programme. • Percentage of targeted supply-side parties that use the programme. • Δ percentage of targeted supply-side parties that can name efficient products and practices (in regard to the targeted end-uses).
Change in knowledge level	• Δ percentage of targeted supply-side parties that can accurately describe energy and performance advantages of efficient products and design practices and their appropriate applications. • Δ percentage of targeted supply-side parties that report reduced perceptions of performance risk for efficient products and practices.
Change in adoption level and promotion of energy-efficient products and practices	• Δ percentage of targeted supply-side parties that report increased promotion of efficient design practices and product specification. • Δ percentage of targeted supply-side parties that report use of efficient design practices and product specification. • Δ percentage of relevant situations in which efficient products or designs are specified.

3.4.3 Development of Baselines

Energy audits. Energy audits usually include detailed data on specific industrial plants or on specific individual building characteristics. Estimates of expected savings for each measure considered by the audit or almost always produced via a modelling process. This information is often used to estimate the (aggregated) savings potential.

Training & Education. As with information centres, participant surveys are most often used to capture information on changes in programme effects indicators. Training programme administrators often administer a quick quiz at the beginning of sessions to assess attendees' knowledge levels. The results of these tests may be used as an indicator of baseline conditions. Similarly, tests conducted at the end of the session could be used as an indicator of programme outputs.

3.4.4 Assessment of Output and Outcome

Verification of Outputs. Verification of programme outputs for education and training programmes generally requires that identification data collected by programme users be processed for comparison to statistics that characterise the targeted population. In the case of programmes that primarily serve residential customers, this is a relatively straightforward task. Each individual user registered with the programme generally corresponds to an element of the target population, i.e. one household or billed customer. However, for programmes that serve the commercial and industrial markets, more effort is required to characterise both the target population and participants. On the supply side, for example, general professional and industrial directories will not identify engineers or equipment suppliers that specialise in serving particular branches. In these cases, evaluators may need to retain the services of an industrial insider to review directories in preparing population lists. Commercial and industrial end-use customers frequently send more than one person to training sessions or information centres. The programmes generally record each of these representatives separately, even though they represent only one unit in the population. In analysing programme outputs, participation needs to be expressed in terms of the number of targeted facilities represented by the participating users. To ensure integrity of the analysis, it may be necessary to match identifying information from the users against directories of firms in the targeted population.

This type of analysis requires that high standards of data quality be maintained in recording identifying information about programme users. Participation data quality control procedures include:

- Determining the basic market function of the participant: end-use customer, engineer/architect, distributor, installer, or maintenance contractor.
- Collecting enterprise characteristics at the time of registration: branch, size (number of employees or energy consumption), location(s), primary business activities (manufacturing, retail, administrative) at various locations.
- Assigning unique identifiers to facilities and/or individuals upon registration.
- Validating facility identification information against customer lists, directories, or other sample frameworks.

- Determining the purpose of each contact. This is particularly important for information centres that provide a variety of services to users.

These procedures must be undertaken at some point when assessing programme participation and estimating market and energy effects. They are most effective and least expensive when they are integrated into programme enrolment and management procedures.

Energy audits have the most direct links to the impacts – energy savings – of the three categories of policy measures. A European study[36] showed that data on input and output indicators are collected in almost all energy audit programmes, but that the outcome and impact indicators are collected less frequently. Nevertheless the costs of collecting data on this later amounts to around 2-4 man-months for smaller schemes and up to 12 man-months for larger schemes. Table 3.7 illustrates characteristics of different options for collecting data on indicators.

Table 3.4.7 Example of different levels of monitoring energy audits

Options	**Coverage**	**Complexity**	**Theoretical cost**		**Information gained from**
			< 100 audits/year	**> 100 audits/year**	
1. Expenditure	All audits	Easily achieved	No extra costs	No extra costs	Application
2. Energy Audit Volumes	All audits	Easily achieved	Negligible extra costs	Minor extra costs – 0.25 man-months/year	Application
3. Savings potential	All audits	More complex - tool necessary, i.e. database	Minor extra costs. Need spreadsheet - 0.5 man-month	Development costs: 6-man months/year. Operation costs: 1 man-month/year.	Audit report
4. Theoretical savings of implemented measures	All audits/ samples	More complex. Need tool and feedback from clients	Operating costs in the range of 2 man-months/year	Operating costs in the range of 4 man-months/year	Questionnaire / site visits
5. Measured savings at site level	All audits/ samples	Complex - need tool, feedback from clients and analytical expertise	Costs in the range of 4 man-months/year	Costs in the range of 1 man-year	Questionnaire (annually)
6. Verified results	Samples	Complex - need tool, feedback from clients on measured data and analytical expertise	Costs in the range of 6 man-months/year (based on representative samples)	Costs in the range of 1 man-year (based on representative samples)	Monitoring on-site level

Assessment of outcomes. Evaluators have generally relied on the following approaches to assess the effects of energy audits and education & training programmes.

- ***Self-reports.*** Most evaluations of education and training programmes make extensive use of participant surveys to develop indicators of programme effects. Because it is difficult in practice to obtain baseline information as part of the programme enrolment process, many evaluations have taken the approach of questioning participants about practices prior to, and after, participation in the programme (in the same

[36] Christensen, W, Aamodt Espegren, K, 2002, *Audit II, Topic Report Monitoring and Evaluation*

questionnaire), which is typically administered within one year of participation.[37] Of course this approach is subject to various kinds of self-reporting bias. One way to mitigate this bias is to use open-ended questions to probe whether customers adopted various practices before or after participating in the programme.

- ***Cross-sectional analysis.*** Unlike general information and labelling programmes, information centre and training programmes generate a set of well-defined and identifiable participants in the course of their operation. Thus, it is possible to assess their effects on knowledge and adoption of efficient designs and products through a cross-sectional comparison of participants and non-participants in the same service territory. For example, a recent evaluation of an information centre supporting the food service industry in California contained surveys of participating and non-participating restaurant owners and managers that supported analysis of the degree to which members of those two groups experienced various kinds of barriers to using energy-efficient products and practices. Non-participants reported much higher levels of performance uncertainty with regard to efficient equipment than the participants. Also, a significantly larger portion of non-participants reported using 'rules of thumb' in purchasing decision making that precluded the selection of efficient equipment.[38]
- ***Historical analysis.*** Generally, education and training programmes reach a relatively small percentage of the target population of end-users and supply-side market parties.[39] Contact between these programmes and their users is also relatively casual compared to that involved in rebate or other economic incentive programmes. For these reasons, historical analyses based on repeated surveys of the targeted populations have not been widely used in evaluations.
- ***Audits and follow-up questionnaires.*** The audit reports contain information on the number of advised measures, energy-savings potential and payback periods. Additional questionnaires or on-site visits are used to research the follow-up of the advice (in a sample). If the questionnaires are repeated annually or biannually, the advised measures are usually distinguished, at least between decided (to implement) and implementation started/finalised. If no information on installation rates is available, a reasonable generic estimate of a typical installation rate is 50% (Sep, 2001, page 1-1).

3.4.5 Assessment of Energy Savings and Emissions reductions

Analysis of Energy Savings. Estimation of energy savings achieved by information centres and training programmes is complicated by a number of circumstances relating to the nature of these programmes and the services they render. Information centre and training programmes are generally used by both facility managers and representatives of supply-side

---

[37] See, for example: XENERGY Inc. 2001. *Evaluation of the Compressed Air Challenge Training Program.* Washington D. C.: US Department of Energy.

[38] Equipoise Consulting, Inc. 1999. *Final Report for Pacific Gas & Electric's 1998 Food Service Technology Center Market Effects Study.* San Francisco: Pacific Gas & Electric Co.

[39] There have been some exceptions to this experience. An evaluation of the Pacific Energy Center found that 60% of the targeted architectural and engineering firms used the various services offered by the Center at least once in the four-year period covered. TecMRKT Works. 1998. *Pacific Energy Center Market Effects Study.* San Francisco: Pacific Gas & Electric Company.

parties who are interested in using efficient products or design techniques in their business or professional practices. The relationship between programme participation and energy savings is very different for end-users versus supply-side representatives. End-use representatives can be associated with specific facilities and questioned regarding the actions they took to increase energy efficiency in those facilities. From there, it is straightforward conceptually, if not practically, to develop estimates of energy savings for a sample of end-users that participated in the programme.

Estimating energy savings from changes in the practices of participating supply-side parties poses much greater challenges. Even an indicative estimate must take into account the number and type of participating supply-side parties; the number and type of facilities they serve in the course of a year; the number and type of energy-efficiency opportunities they encounter in those facilities; their practices prior to and after the training sessions in which they participated. It is difficult for most contractors, vendors, and engineering professionals to report accurately on any one of these factors, much less all of them at once.

Savings from participating end users. Analysts interested in estimating energy savings achieved by end-use facility owners and managers who participate in information centre and training programmes face a number of challenges, including:

- ***Heterogeneity of customer facilities.*** All education and training programmes serving the commercial and industrial sectors deal with customers that vary significantly in terms of size, industrial branch, and energy efficiency opportunities available. In estimating savings for these kinds of programmes, it is not sufficient simply to multiply the number of measure adoptions attributable to the programme by a uniform factor. Rather, the specific characteristics of at least a sample of participating customer facilities should be captured and taken into account in savings estimates. The evaluations cited above used surveys to capture basic information about sample facilities, such as number of employees, size, equipment inventories, hours of operation, and basic functions. Then, using secondary information on energy use indices or simple engineering formulae, they estimated energy use in the systems or building elements addressed by the programme.
- ***Heterogeneity of potential energy efficiency measures.*** By their nature, information centres and training programmes are designed to support customer and vendor implementation of the full range of efficiency measures appropriate for the targeted end-uses and facilities. For example, the Compressed Air Challenge programme discussed above promoted roughly 20 measures that addressed four different elements of compressed air systems. To estimate energy savings in this situation, evaluators must develop energy savings calculations or factors for the measures that participants frequently report undertaking. In the case of relatively homogeneous end uses, such as residential lighting, simple unit estimates can be used. However, for more complex kinds of measures, evaluators have typically developed savings factors based on industry experience which are then applied to estimates of energy consumption in sample facilities. For well-established technologies, factors may be derived from other evaluations. In the case of the compressed air systems programme, for which documented energy savings experience was sparse, a panel of industry experts was convened to develop energy savings factors for the most common measures undertaken, and for various combinations of those measures.

Given that participants in education and training programmes can be identified, it should be possible to use more data-intensive methods, such as on-site observation, measurement, and bill analysis to estimate energy savings. However, evaluators will need to balance the precision in measuring changes in energy use versus cost. Moreover, the relationship between participation and changes in energy consumption are much more tenuous in the case of information programmes versus incentive programmes. Thus, greater precision in estimating changes in energy use may not coincide with greater confidence in the overall estimate of energy savings attributable to the programme. Note that the authors have not encountered any instances in the literature in which direct observation or bill analysis techniques were used to estimate the energy savings associated with information centre or training programmes.

Measured savings. Savings can be presented in several ways. Firstly, the theoretical savings: based on data concerning measures that have actually been undertaken and the calculated savings in the audits, it is possible to estimate the energy savings. This will be a rough estimate, since there have been no physical measurements. Secondly, the measured savings (at site level): sometimes it is possible to follow the annual changes in energy use on the same level as the measures implemented. Changes in the overall annual energy use are often made. In general, audit-predicted savings tend to be at least somewhat higher than measured savings. The Sep Metrics Handbook advises adjustments for realisation rates (measured savings to audit-predicted savings) of 60% for residential audits and 90% for commercial, industrial and agricultural audits (Sep, 2001 page 1-2). Thirdly, verified results: all physical measures will be individually monitored and verified by a third party. This verification is often undertaken by using sampling methods.

The Danish case example on the free audit (included in Volume II) reports that annual impacts from audit activities are approximately 100 GWh (1st year electricity savings). However, measurements are not always possible. The Danish electricity network companies offer free-of-charge energy audits for non-residential customers. The audit reports contain information on annual electricity consumption and suggest a number of measures that can be implemented in order to reduce electricity consumption. The suggestions include estimates of investment costs and potential electricity savings. While the number of suggestions that are actually implemented can be established fairly easily, it is frequently impossible to verify the actual electricity savings without installing extra meters. Customer investment costs are often embedded in other costs, for example as part of a general overhaul of a production unit, which can at times also be difficult to establish. As an alternative to measured electricity savings and investment cost, the Danish electricity network companies apply the estimates made in the audit reports, plus feedback on the degree of implementation of suggested improvements, in order to arrive at figures for achieved electricity savings and socioeconomic costs.

3.4.6 Calculation of Costs, Cost-efficiency and Cost-effectiveness

Costs relating to information centres and training programmes concentrate on cost-efficiency. This ratio between the input and output provides information on the management of the programme. Energy audits themselves contain information on the cost-efficiency of advised measures, while the potential savings versus programme costs can be used as an indicator of efficiency.

A recent evaluation of the Danish free-of-charge audit programme offered by the electricity network companies operates with various cost perspectives (i.e. utility, society and client), as well as two differing timeframes, namely first year and 15 years (summarised in Dyhr-Mikkelsen, 2005). For example, the electricity network company costs to achieve an impact of 112 GWh first-year savings in 2002 was 9 eurocent/kWh (utility perspective). The CO_2 shadow price (society perspective) of the planned audit activity for 2003-2004 was expected to be between -2.4 EUR/ton CO_2 (for industry larger than 500 MWh/year) and 41.4 EUR/ton CO_2 (for agriculture below 100 MWh/year) depending on the client group in question and assuming a lifetime of 15 years for the implemented changes. A case study of 10 audited companies, which was carried out as part of the evaluation, showed a CO_2 shadow price ranging from -35 to +84 EUR/ton CO_2, assuming a lifetime of 15 years and only counting electricity savings. The general government guideline for energy efficiency activities is 19 EUR/ton CO_2. However, the 10 clients all benefited from the audit, as can be seen in Table 3.12 (client perspective). The case study showed – as does the documentation prepared by the electricity network companies – that the costs and benefits vary greatly depending on the type of enterprise. Table 3.8 includes additional information.

Table 3.8 Danish free-audit programme and costs

	Client perspective (EUR)			CO_2 shadow price (EUR/ton CO_2)			
Case enterprises by branch and size (MWh/year)	Change in electricity bill	Investment minus subsidy	Net savings	NPV	Interest rate +/- 2%point	Electricity price +20%	5 years lifetime
Agriculture 100-500	-4,916	2,346	2,570	22.4	27.7/17.2	**14.5**	68.9
Industry >500 (a)	Not calculated						
Industry >500 (b)	-68,699	320	68,379	**-34.6**	**-37.1/-38.0**	**-45.6**	**-32.7**
Industry 100-500 (a)	-15,893	9,662	6,231	41.1	47.8/34.4	33.2	100.5
Industry 100-500 (b)	-38,203	6,340	31,863	**-3.6**	**-0.4/-6.7**	**-11.6**	24.9
Trade & service >500	-20,440	10,765	9,675	22.0	27.3/16.9	**14.1**	68.4
Trade & service 100-500	Not calculated						
Trade & service 20-100	-4,544	928	3,616	**6.3**	**10.2/2.5**	**-1.5**	40.9
Public >500	-140,226	7,646	132,580	**-24.6**	**-23.1/-26.1**	**-32.5**	**-10.7**
Public 100-500	-12,496	4,215	8,281	83.7	94.1/73.5	75.5	173.9
Public 100-500 incl. heating	*-21,519*	*4,215*	*17,304*	***10.0***	***18.0/1.9***	-	*81.6*

Source: Dyhr-Mikkelsen, 2005

3.4.7 Levels of Evaluation Effort

As with more general types of information policy measures, tracking outputs from education and training programmes is a necessary and often difficult first step in the evaluation effort. Procedures to capture information about users in the normal course of programme operation are absolutely essential here. Experience has shown that *post hoc* reconstruction of participation information is time consuming, inexact, and expensive. Comprehensive evaluations will require the development of a credible method for estimating energy savings based on primary research on a sample of participants. Moreover, information on the knowledge and technology adoption patterns of non-participants will be useful in developing assessments of the net impacts of the programme. Targeted evaluations will generally consist of fewer surveys or may substitute lower cost data collection mechanisms such as focus groups for full-scale sample surveys. See Table 3.9.

The country case examples as included in volume II range from level A (free-of charge electricity audit, Denmark) to level C (energy audits in buildings, Korea).

Table 3.9 Evaluation activities associated with different levels of effort: energy audits and education & training programmes

Level A **Comprehensive Evaluation**	**Level B** **Targeted Evaluation**	**Level C** **Programme Review**
Development of policy measure theory and specification of indicators		
Literature review. Expert interviews. Consumer focus groups.	Literature review. Expert interviews.	Literature review.
Characterisation of activities		
Early development and analysis of programme participation records. Corroboration of records with programme staff and contractors.	Early development and analysis of programme participation records. Corroboration of records with programme staff and contractors.	Early development and analysis of programme participation records. Corroboration of records with programme staff and contractors.
Estimation of changes in indicators		
Population-based sample surveys: participant versus non-participant.	Population-based sample surveys: participants only.	Small sample surveys of participants.
Baseline development/estimation of net outcomes		
Cross-sectional analysis. Survey of participants versus non-participants. *Self-reports.* Customer and supply-side party surveys covering self-reports of programme effects. *Expert interviews:* Corroboration through programme staff, contractors, related vendors.	*Self-reports.* Customer and supply-side party surveys covering self-reports of programme effects. *Expert interviews:* Corroboration through programme staff, contractors, related vendors.	*Self-reports.* Customer and supply-side party surveys with small samples covering self-reports of programme effects. *Expert interviews:* Corroboration through programme staff, contractors, related vendors.
Estimates of energy savings		
Engineering analysis. Energy measuring for implemented measures	Engineering analysis. Annual energy use data.	Engineering analysis. Energy savings measures models.

3.5 Conclusions

Information policy measures and programmes cover a wide range: from general information campaigns and information centres to labelling, education and training, and energy audits. Most hypotheses dealing with general information use some kind of causal model whereby knowledge, attitudes and behaviour are the main features.

Sometimes there are no baselines included in the evaluation, especially for information campaigns, while this could be done. This chapter includes suggestions, not only for baselines, but also for outcome indicators.

For energy audits the impact in energy savings is a clear objective and (at least the potential) energy savings are included in evaluations. Following the Sep Metric Handbook we recommend paying special attention in the evaluation to the adjustments for realisation rates from audits.

While Chapter 1 discussed several options for looking at costs for information policy and measures, we concluded that costs mostly relate to the ratio between input and outputs, i.e. cost-efficiency.

4. EVALUATION OF ECONOMIC INCENTIVES POLICY MEASURES AND PROGRAMMES

4.1 Introduction

The economic incentives policy measures and programmes addressed in this *Guidebook* include a broad range of subsidies, rebates, taxation, grants, loans etc. The authors also hold the opinion that the policies dealing with bulk purchasing, technology procurement and certificate trading systems should also be included in this category. Case examples for subsidies, rebates and tax credits from Denmark, Korea and the Netherlands are used to illustrate this theory (see Table 4.1). A detailed description of these cases, plus the Italian case example on certificate trading, can be found in Volume II of the *Guidebook*. Information from these case examples is also used to illustrate elements such as output indicators, development of baselines and energy savings.

Table 4.1 Subcategories and case examples for regulation policy measures and programmes

Subcategories	Case examples	Country
Project or product-related subsidies (rebates)	Rebate programme for highly efficient electric inverters	Korea
	Energy premium scheme households	Netherlands
	Financial incentives for DSM	Korea
Reduced-interest loans		
Financing guarantees		
Third-party financing facilitation		
Targeted taxes, tax exemption, tax credits	Energy Investment Reduction (EIA and EINP)	Netherlands
Bulk purchasing		
Grants		
Technology procurement		
Certificate trading systems	Criteria adopted for evaluating primary energy savings in end-users / EE Certificates	Italy

This chapter starts with a section on the objectives and main kinds of economic incentive policy measures and programmes. The aforementioned subcategories are grouped in Table 4.1 into the following:

1. Price-reducing policy measures and programmes.
2. Taxation systems.
3. Policy measures and programmes dealing with financing arragements.
4. Ensuring a minimum market share.
5. Certificate trading systems.

Sections 4.3 to 4.6 analyse the first four of these groups according to the seven key elements discussed in Chapter 1:

Element 1: Policy measure theory used
Element 2: Specification of indicators for the success of a measure
Element 3: The baselines for the selected indicators
Element 4: Assessment of outputs and outcomes
Element 5: Assessment of energy savings and emissions reductions and other relevant impacts
Element 6: The calculation of costs, cost-efficiency and cost-effectiveness
Element 7: The level of evaluation effort

Based on the case examples in Volume II and the discussions at the experts meeting, we see the following issues of considerable interest for evaluating policy measures and programmes dealing with economic incentives:

- In the baseline development special attention should be given to the question 'How to handle free-riders and spillover effects?' These two elements may have considerable impact on the conclusion whether the policy measured did have a (major) impact or not.
- It is difficult to find the appropriate level of subsidies. Evaluations should pay attention to this in their work, as well as in the reports to improve the learning process in policy measure development and implementation.
- Elements are often included in the hypothesis regarding the assumptions on expected market changes and relations with lowering future prices for products. This is one of the more important elements in the outcomes of these measures, and these assumptions should be researched not only in the evaluation shortly after the policy has come to an end, but also in the long run.
- A time lag should be included in the discussion regarding the level of evaluation effort. There is a trade off between the timing of evaluation at the end, or shortly after a policy measure has come to an end, and the related increase in the market share that will be realised some years later.
- Economic incentives are increasingly no longer used as a stand-alone policy measure, but combined with others, e.g. with Voluntary Agreements. This requires that some additional points be taken into account in the evaluation. Chapter 6 discusses this in more detail.

The authors' general feeling is that economic incentives policy and measures should be a priority area for evaluation. They span a wide range of policies and the government is always an important player and financier. In addition, the main theories behind these policies are economic: changing prices. So evaluating cost-efficiency and market changes in particular should be improved.

4.2 Objectives and Main Types of Economic Incentives Policy Measures and Programmes

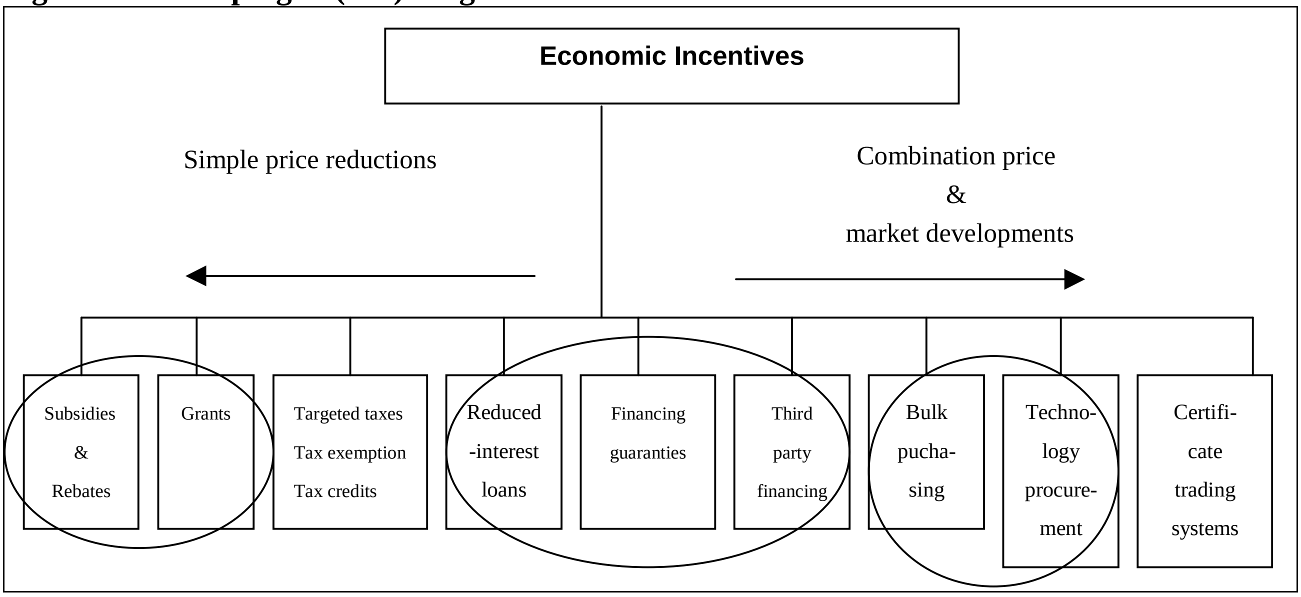
Economic policy measures offer the stakeholders financial incentives to overcome financial difficulties at the consumer level and to foster individual decisions with regard to energy-efficient devices or adopt specified energy-efficient technologies in business processes, e.g. replacing equipment, remodelling, and new construction projects.

- ***Project or product-related subsidies (rebates).*** Rebates are offered for the documented use of specific products or construction techniques. Rebates are generally gauged according to the efficiency level and quantity of equipment installed.
- ***Grants.*** An amount of money given to an individual or organisation for a particular purpose.
- ***Targeted taxes, tax exemptions, and tax credits.*** Several European countries offer tax credits or accelerated depreciation for purchasing specified energy-efficient equipment. In some countries, partial exemption from fuel taxation is offered to facilities that meet agreed requirements for voluntary energy use reduction.
- ***Financing guarantees.*** Programme sponsors may offer credit guarantees to reduce risk premiums charged on loans to finance energy efficiency projects.
- ***Third-party financing facilitation.*** Third-party financing approaches, such as energy performance contracting, are used to finance energy efficiency projects. They often include a subsidy or credit guarantee that reduces the cost of the project to the customer.
- ***Reduced-interest loans.*** Some organisations offer reduced-interest loans to finance projects that incorporate specified energy-efficient technologies.
- ***Bulk purchasing.*** Organisations may aggregate large orders of energy-efficient equipment to receive favourable pricing from manufacturers. These price reductions are then passed on to the final customers purchasing the equipment.
- ***Technology procurement.*** A process through which a commodity, service or system is procured, and for which development of new technical solutions is essential in order to meet a specified requirement by a buyer (or group of buyers). The development work may concern the product, system or the production process for which it is developed.
- ***Certificate trading systems.*** A system of green (or white) energy certificates is used to facilitate the market for renewable energy, energy savings or for energy efficiency improvements.

Though presented in this *Guidebook* as specific instruments, in reality these economic incentives are often combined with other policy measures. For example, promotion: retailers participating in energy efficiency information or rebate programmes using forms of advertising to leverage the effect of the programme to their advantage and thus sell more high efficiency products (Violette, 1995).

These mechanisms are used in many countries, but the level of the incentives may be very different. When evaluating these incentives one should keep in mind that individual national situations are very different.

Figure 4.1. Grouping of (sub)categories of economic incentives



To facilitate the following exposition, the authors have organised the policy subcategories into four groups. For the fifth group (certificate trading systems) we refer to the Italian case example in Volume II. Certificate trading systems are a rather recent development and evaluations on this subject are currently not known.

The first group consists of project or product-related subsidies (rebates) and grants. All measures focus on simple and immediate price reduction, thus enhancing higher sales volumes. The second group contains all types of taxes advantages and disadvantages, including fiscal rules for investment periods that benefit energy-saving technologies. The group on 'financing arrangements' always includes action of additional (market) organisations willing to provide reduced-interest loans, financing guarantees, or a third-party financing facilitation. All contain a similar mechanism: financing of energy efficiency measures using special loans that have to paid back based on energy savings. The fourth group contains policies that combine price with ensuring a minimum market share, either by price reduction through bulk purchases or a take-off market for new products. The next section presents the main elements in greater detail than the later sections, to avoid repetition. For a good overview of interesting suggestions on evaluation items we advise readers to study Section 4.4, and later in combination with Section 4.3 below.

4.3 Price-reducing Policy Measures and Programmes

4.3.1 Policy Measure Theory

The policy measure theory statement provides the basic framework for the evaluation as it identifies the relevant market parties (the policy measure domain) and the hypotheses effects.
The application of economic incentives is based on the assumption that the problems for energy savings and the environment are partially due to a failure of the market system.

Specification of Policy Measure Domain

The groups of market parties who are likely to be affected by the programme are: end-users, intermediate organisations (e.g. vendors and installers) and manufacturers. A subsidy for a manufacturer is either a subsidy for new or improved appliances etc. that they wish to produce and sell, or for energy savings by energy users. In the first case, this subsidy is for R&D efforts and we do not include evaluation on R&D in this *Guidebook* (see Chapter 1), and in the second the producer is seen as an end-user.

The main domain of subsidies and rebate programmes includes:

- Consumers or households: to encourage energy efficiency investments in implementing energy efficiency measures by end-users, e.g. in existing buildings and equipment, or the adoption of energy-efficient techniques.
- Industrial companies: to encourage investment in energy-efficient appliances, including process technologies.
- Governmental organisations and the service sector (commercial and non-commercial): to encourage energy efficiency investments in implementing energy efficiency appliances (heating and cooling, electrical equipment) and measures in existing buildings, as well as organisational changes to improve the energy management.

To avoid free-riders or insufficient use, one should pay suitable attention to targeting the programme towards the selected groups and to the mechanisms for informing these groups. Hypotheses on this subject are well known from information programmes, as included in Chapter 3.

Statement of Policy Measure Effects Hypotheses

The hypotheses on the effects of subsidies and price rebate programmes are based on encouraging investments through reducing (perceived) costs. The subsidies or grants help customers to implement energy efficiency measures, in the first place by reducing costs. Often the costs of the more energy-efficient solution are still higher than traditional alternatives. For this reason information on the money saved through reduced energy consumption is given to the consumer.

Subsidies should be viewed as a temporary measure to mobilise consumers, to prepare for new regulations, or to promote energy-efficient technologies by creating a larger market than would otherwise exist.

The objective is to use the available funds to create a sustainable market transformation, when reducing perceived risks to market players, capturing the attention of otherwise apathetic or uninformed customers, and temporarily reduce prices until market trends naturally force these prices downward (Gibbs and Townend, 2000). So, in addition to the financial hypotheses, other hypotheses are often included for:

1. Reducing the perception of performance risk: the consumer feels that a bad product would not be subsidised.
2. Reducing information search time/costs: product- and measurement specifications are easily available on the subsidy or grant list.
3. Decrease the risk to manufacturers of introducing new energy-efficient product lines; the subsidy scheme will generate an additional market.
4. Reducing the retailers’ risk in stocking and displaying energy-efficient products.

4.3.2 Specification of Indicators

Input indicators are monies used for subsidies or grants and monies used for promoting the scheme as well as man-hour capacity.

Output indicators are the countable products or implemented measures such as the:
- number of subsidised projects;
- percentage of eligible facilities that participate in the programme;
- market share of qualifying products;
- number of subsidised agencies;
- ratio investment / related investment costs;
- ratio requested/realised subsidy.

Examples of **Outcome** indicators include awareness of energy-efficient products, positive attitudes towards the new energy efficiency products and changes in square meters for the display of energy-efficient products. Also the proper use of the product is important as outcome indicator, as it has a major influence on the value of impact indicators.

The **impact** indicators are (calculated) energy savings and CO_2 reduction related to the subsidised products and measures. As this policy measure also intends to change the market one can also include the change in market share percentage or the point where the energy-efficient product becomes competitive. If the subsidy is targeted towards demonstration projects then the replication rate of the subsidised technology in the market (when the programme is finalised) could also be used.

4.3.3 Development of Baselines

Subsidies and policy rebate programmes assume that, when granting economic incentives, at some point the changes in implementation level, consumer and manufacturer behaviour etc. will result in energy savings.
The main elements for a baseline study include:
- Size and composition of target markets.
- Pre-programme awareness or knowledge levels.
- Pre-programme information and education sources.
- Extent of exposure to (and use of) pre-programme education or information sources.
- Pre-programme status of the target market relative to the intended results of the programmes.
- Pre-programme adoption patterns.

In general, no (or only a few) examples of baseline examples for these economic incentives have been found. As a result the outcome and impact for subsidies and grants are generally overestimated.

4.3.4 Assessment of Output and Outcome

Verification of Output

The organisation managing a price-reducing measure almost always has to report to the government on how the money was spent. A management report will be generally be presented and discussed on a regular basis (monthly, quarterly, biannually, annually). The administration of the subsidy scheme contains most of the information required for the output indicators. This information is therefore easily available to the evaluator and is often of a high quality. The indicator most often discussed concerns the realised management costs, especially when the realised number of subsided projects, grants etc. is (much) lower than expected.

The output indicators dealing with market shares need additional market information. This can sometimes be obtained from publically available statistics, but it often has to be bought from market research companies and sometimes specially researched for the programme. In the latter case, the research is rarely conducted during the programme implementation, but is more often conducted at the end as part of an ex-post evaluation.

Consumers who could use the subsidy and were targeted by the scheme (e.g. small to medium-sized industries, low-income households etc.) often do not take advantage of it simply because they are unaware of its existence. This is often due to the difficulty of adequately publicising the existence of these incentives to the multitude of consumers concerned. Consumers may also be dissuaded from applying, as procedures to obtain the grants are often too bureaucratic, requiring complex forms to be completed and entailing long delays in obtaining approval. An ex-post evaluation could include this in its assessment of the output indicator number of subsidies or grants.

Verification of Outcomes

As outcome indicators are often related to awareness, attitudes, usage and changes in habits, the verification of outcomes contains lots of similarities to that of information policy (see Chapter 3). To summarise:

- Random-sample surveys.
- Multi-client surveys.
- Retail location inventories.
- Self-reports.

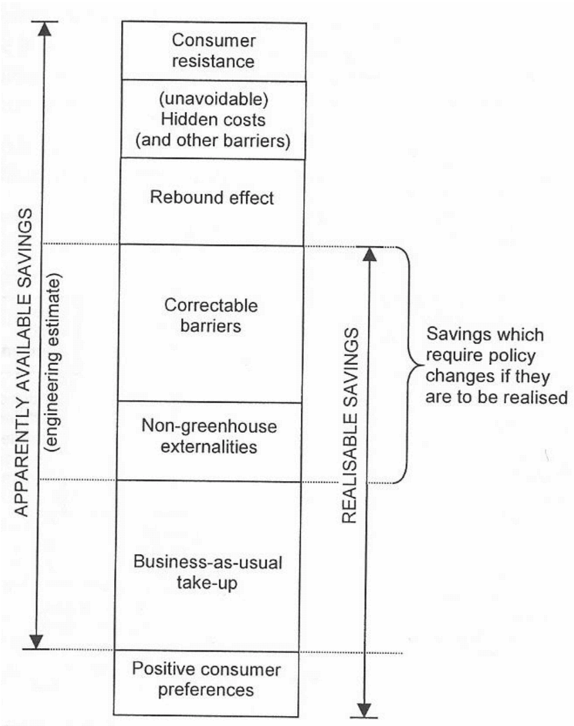
4.3.5 Assessment of Energy Savings and Emissions Reductions

Programmes that provide financial support for the installation of energy efficiency measures have an assumed direct impact on energy efficiency. The mechanism by which the programme operates is simple and well understood and the programme beneficiaries are easily identifiable. Therefore most such programmes are 'measured' in terms of calculated energy savings and/or carbon savings. This is also the case where financial support mechanisms form a part of other programmes, e.g. market transformation programmes that include a subsidy for sales or targeted products.

In the majority of the monitoring and evaluation reports the calculated energy savings are based on the number of subsidised measures, plus average usage time and calculated energy savings based on standard situations (e.g. an energy-saving bulb versus a standard bulb). In addition, subsidies in industrial companies or buildings often use the energy use in previous periods (as reported in the subsidy application form) as a reference. Examples of these are available in the IEA-DSM INDEEP database (http://dsm.iea.org) and are presented in the INDEEP analysis report 2004.

An assessment of energy savings and potential/realisable gains is seldom conducted (see Figure 4.2. for an example of this). Sometimes attention is given to free-riders (consumers who would have carried out the investments even without the incentive) and to the rebound effect, ('inefficient' use of the new appliance) both resulting in lower energy savings relating to a subsidy programme as calculated in a standard way.

Figure 4.2 Energy savings, potential and realisable gains



Source: Grubb et al., 1993 page 410, as presented in Ostertage, 2003 page 3

Elements that could be taken into account for adjusting the calculated energy savings include:
- Rebound-effect: subsidies or price rebates result in lower (variable) energy costs, that may eventually contribute to 'inefficient' energy use. A reduction in the marginal cost of energy services causes reduced operational costs or lower prices of energy-efficient products. An increase in total energy consumption might result from the fact that consumers can purchase new, additional appliances sooner (or keep the old appliances running) in addition to using the new one. In some programmes studying the subsidy on investments in energy efficiency reveals that even utilities, which invested heavily in energy-efficient technologies, experienced no reduction in electricity demand compared with utilities that largely avoided such subsidies.
- Free-riders: consumers who would have bought energy-efficient devices with or without financial incentives. Subsidies have sometimes paid consumers to do what they would

otherwise have done of their own accord. This 'free-rider' effect contributed to the Norwegian government scrapping its grant programme in 1993.
- Attention-effect: the number of consumers who are granted subsidies due to the attention paid to the existence of the energy-efficient technique. This is a correction of the free-rider effect. The technique is particularly relevant for subsidy receivers, but they were not aware of the scheme.
- Baumol-effect: correction for income effects caused by lower energy prices. Reduced energy prices might result in higher (energy) consumption.

The calculated emission savings are always based on key figures for emissions relating to a specific fuel or to electricity in general. Table 4.2 contains an example of key figures used in the Dutch MAP subsidy programme.

Table 4.2 Some key figures in the Dutch MAP subsidy programme 1990-2000

Measure	Units	Key figure CO_2 reduction
Wall insulation	51 m3 a house	13 Kg/m2
Roof insulation	56 m3 a house	11 Kg/m2
HR boiler in new houses	1	212.4 Kg
HR boiler replacement	1	319 Kg

Source: Beerenschot 2001 page 77/78

4.3.6 Calculation of Costs, Cost-efficiency and Cost-effectiveness

Cost (effectiveness) can be studied from the following perspectives:
- the consumer/investor;
- the government; or
- at a national level, taking into account a trade-off between economic sectors and energy savings options.

Four economic indicators are in use to evaluate, from the investor's point of view, the profitability of an investment that is influenced by the subsidy or grant. These indicators are:
- The cost of energy saved: the investment cost dived by the expected annual energy saving produced by this investment for the (economic) lifetime. Values that are often used for lifetimes in subsidy schemes are 10 years (for an industrial project) and up to 20 years (for domestic and service sector projects).
- The gross payback period of the investment: the period after which the investment is compensated from the energy savings induced by this investment.
- The cost:benefit ratio of the investment: the ratio between the amount invested in the subsidised or granted measure or appliance and the annual gain over the lifetime.
- The internal rate of return on investment: the discount rate at which the net present value of the benefits is equal to the discounted cost of the investment.

From the governmental perspective, investments and costs (specified in subsidies and administrative costs) have to be evaluated and compared to benefits. These are generally included in the input and output indicators. We recommend that evaluations should pay additional attention to the cost effectiveness.

If the simple ratio subsidy:cost effectivity is restricted to the individual subsidy, this might result in a too optimistic ratio, especially if the programme is targeted towards industrial or commercial organisations. In the evaluation conducted on three subsidies in the Netherlands (CEA, 2002) the evaluator concluded that:

- The expected costs effectivity of € 4.81 per ton CO_2 is relatively very positive.
- The overall stimulation of several governmental programmes is much higher and thus too the overall governmental costs; additionally, the evaluator considered that the specific subsidy alone was too small to influence investment decisions and that only the combination of several governmental subsidies could achieve this.
- While combined governmental support is estimated to be around € 50 per ton CO_2, the total costs for the energy users could go up to € 80 per ton CO_2.

However, the cost-effectiveness for the household could also be very different, depending on what is included in the calculation. An analysis of the multi-year Dutch Energy Premium Scheme, which is included as a case example in Volume II, (Boonekamp, 2000) contains two types:

- Ongoing cost-effectiveness that takes into account the annual subsidy (as cost) and the cumulative savings (as investments in previous years continue to save energy).
- Cumulative cost-effectiveness that sums up all subsidies and energy savings up to the end of a period.

Table 4.3 illustrates the various figures that result from using different types of definitions for cost-effectiveness.

Table 4.3. Overview of calculated (scenario) effects for the Dutch Energy Premium scheme on energy use in households, for the years 2005 and 2010

	2005	2010
Energy savings (PJ)		
Gas	3.1-13.0	5.6-20.5
Electricity	1.5	2.5
CO_2 emission reduction (Mton)		
Insulation	0.18 – 0.74	0.31 – 1.17
Appliances	0.15	0.25
Subsidy (mln € a year)		
Insulation	71 – 89	70 - 95
Appliances	67	84
Total	138 – 156	154 - 179
Ongoing effectiveness (€/ton CO_2)		
Insulation	414 – 122	221 - 81
Appliances	437	333
Average	428 - 176	270 – 126
Cumulative effectiveness (€/ton CO_2)		
Insulation	275 – 81	293 - 104
Appliances	293	401
Average	284 - 117	338 - 158

Note: the two figures for insulation are for minimum and maximum implementation of policies
Source: Boonekamp 2000 page 5

4.3.7 Levels of Evaluation Effort

All evaluations will use the (management) reports that are produced during the implementation of a subsidy or grant programme. However, energy savings and estimates of costing additional efforts could be included in a more comprehensive evaluation.

Table 4.4 Evaluation activities associated with different levels of effort: subsidies and price rebate policy measures and programmes

Level A Comprehensive Evaluation	Level B Targeted Evaluation	Level C Programme Review
Development of programme theory and specification of indicators		
Literature review. Management reports on subsidy and grant programmes Expert interviews.	Literature review. Expert interviews	Literature review.
Characterisation of programme activity		
Quality control of programme records, Analysis of programme records Programme management interviews Analysis of information activities	Analysis of programme records Management reports on subsidy and grant programmes Programme management interviews	Analysis of programme records Management reports on subsidy and grant programmes
Estimation of changes in indicators		
Analysis of programme records and management reports Sample surveys on awareness of and barriers to make use of the price reducing programmes (annually).	Analysis of programme records	Analysis of programme records
Baseline development/estimation of net impacts		
Develop a baseline in the evaluation and include free-riders and market segmentation	Develop a simple baseline in the evaluation	
Estimates of energy savings and emission reductions		
Include a range in the calculated energy savings taking into account selected adjustments such as free-riders or rebound.	Include usage changes in the calculation	Calculation based on key figures
Estimates of costs		
Cost from different perspectives Use of cost-efficiency indicators Include cost-effectiveness	Costs from more than one perspective (e.g. government and consumer) Use of cost-efficiency indicators	Costs from one perspective (e.g. just government)

4.4 Taxation Systems: Targeted Taxes, Tax Exemption and Tax Credits

Since the early 1990s tax policies and measures have increased considerably in all IEA Member Countries. In 1999, 65 tax instruments to encourage GHG (greenhouse gas) emission reductions were in force. Most of these were not directly linked to GHG emissions, but affected emissions indirectly. Table 4.5 illustrates the wide variation of tax policies and measures.

Table 4.5 Tax policies classification by country, 1999

Sector	Political Declaration	Taxes			Tax Relief		Other	Total
		Product Taxes	Emissions Taxes	Technology	Tax Reduction	Tax Credit		
Residential/ Commercial		2			2	3		7
Transport	1	9	2	6	2	6	2	28
Industry/ Manufacturing		3	1		5	3		12
Power Generation		3	1		5	3		12
All	4	4	7	1		4		20
N/A						1		1
Total	**5**	**19**	**12**	**7**	**14**	**21**	**3**	**81**

Source: IEA, 2000 page 26

4.4.1 Policy Measure Theory

Policy Measure Domain

Targeted taxes are compulsory payments to the government. The goal of energy, carbon tax policies etc. is to correct market failures by internalising economic externalities, enabling the price of goods and services to reflect full social and environmental costs. The taxes make the greatest sense economically and environmentally because they directly tax the externality – inefficient energy use and/or carbon emissions. These taxes withdraw financial resources from the target group where it exhibits undesirable behaviour on energy use. The environmental or energy taxes on fuel and electricity use are the most well known. Table 4.6 illustrates taxes on transport fuels and energy products for industrial use. In several European countries (e.g. Denmark and the Netherlands) an energy tax for non-industrial use was also introduced in the late 1990s.

Table 4.6 Taxes on energy products in some selected countries, in purchasing power parity, 2000

Country	Petrol (gasoline) unleaded		Diesel		Diesel/Gas Oil *(industrial use)*		Coal *(industrial use)*		Natural Gas *(industrial use)*	
	\$PPP / 1000 litre	\$PPP / ton CO_2	\$PPP / 1000 litre	\$PPP / ton CO_2	\$PPP / 1000 litre	\$PPP / ton CO_2	\$PPP / 1000 kg	\$PPP / ton CO_2	\$PPP / 1000 m^3	\$PPP / ton CO_2
Denmark	395	164	272	95	206	*72*	163	*67*	28	*15*
Finland	558	232	324	113	55	19	33	14	28	15
France	590	245	370	129	78	27	0	0	1	1
Germany	495	205	313	109	40	14	0	0	33	17
Netherlands	583	242	336	117	102	36	11	5	55	29
Norway	520	216	403	140	46	*16*	46	19	93	49
Spain	490	203	356	124	104	36	0	0	8	4
Sweden	456	189	295	103	183	*64*	126	*52*	105	*56*
Switzerland	356	148	372	129	1	1	0	0	0	0
UK	630	261	645	224	40	14	0	0	0	0
USA	101	42	116	40	na	na	na	na	na	na
Japan	320	133	124	43	4	1	na	na	23	12

Source: Baranzini, Goldenberg and Speck, 2000.

Tax credits, tax reductions and accelerated depreciation are fiscal incentives that **reduce** the tax paid by consumers who invest in energy efficiency. This policy instrument is generally targeted towards commercial and non-commercial organisations. As it also often requires specific provision in the relevant legislation, the target groups are either very broad (e.g. all industrial companies) or very specific (HE air conditioners).

Policy Measure Hypotheses

Taxes may discourage the use of energy sources, the purchase and/or use of less energy-efficient products. The hypotheses are that the consumer responds to a price change. An additional assumption could be that the response also depends on the source of the price change. In that case the basic question is: do changes in the consumer price resulting from the introduction or modification of environmental taxes send a different signal to the consumer than when changes in the consumer price result only from a manufacturer's price change.

The hypotheses on the effects of tax reductions etc. are in line with those for subsidies and price rebates: encouraging investments by reducing costs. Since these schemes are incorporated into the taxation system it is assumed that there are fewer problems in informing target groups of these costs reduction opportunities.

Taxes may vary according to indices representative of the energy use (or efficiency) of a product. A tax reduction may also decline over time, as the logic behind the measure is that the reduction is a temporary one not only targeted towards increasing the investment in energy-efficient products, but also towards reducing the price of these products and to changing the market over time.

4.4.2 Specification of Indicators

The majority of the indicators are the same as for subsidies and grants (see relevant sections). **Input** indicators are monies involved in the taxes and monies used for providing information to the customers as well as man-hour capacity.

Output indicators are different for the targeted taxes and the tax reductions. The output of the targeted taxes refers to the amount of money reuslting from these taxes. For the targeted taxes the outputs are more countable products such as the:
- number of projects with tax reduction etc.;
- amount of tax reduction;
- ratio investment / tax reduction.

Examples of **Outcome** indicators include awareness levels of energy-efficient products, positive attitudes towards new energy efficiency products and changes in future decision-making processes.

The evaluation of the Dutch case example on the Energy Investment Scheme, included in Volume II, showed that the only around 40% of the participants use some kind of investment criterion (de Beer 2000). For this group, the criterion the company used can be compared to the payback period of the measure implemented and the evaluators concluded that:

- 68% (of these 40%) should be assigned as free-riders (a questionnaire confirmed that 51% of the respondents call themselves free-riders.
- 16% (of these 40%) found that the tax reduction measure was profitable, while an additional 17% considered the investment profitable even without the tax reduction.

The **impact** indicators are (calculated) energy savings and CO_2 reduction relating to the taxes (i.e. products, measures and investments). As this policy measure also intends to change the market one can also include the change in market share percentage or the point where the energy-efficient product becomes competitive.

4.4.3 Development of Baselines

Tax-related measures assume that reducing the investment costs will result in changes in implementation level, consumer and producer behaviour etc. resulting in energy savings.
A baseline study is a business-as-usual scenario including elements such as:

- Size and composition of target markets.
- Investment payback periods.
- Pre-programme adoption patterns.
- Energy prices.
- Economic growth.

4.4.4 Assessment of Output and Outcome

The tax offices or the organisation that manages the taxes reductions almost always has to report to the government on the amount and way in which these targeted taxes are collected or the tax reduction is implemented. Governments often use these reports to inform the national parliament, which sometimes result in parliamentary debates. The notes on these debates are useful for the evaluator as well as the regular management reports.

Evaluators could also consider the unintended effects that might be caused by targeted taxes. For example, high taxes can deter consumers from buying new products, thus slowing down the introduction of new technologies or concentrate markets on the most affluent sections of the population.

4.4.5 Assessment of Energy Savings and Emissions Reductions

In the majority of the monitoring and evaluation reports the calculated energy savings are based on the number of taxed or tax-reduced measures and calculated energy savings based on standard situations. Emissions reductions are also normally calculated using simple key figures. For more details see Section 4.3.5.

4.4.6 Calculation of Costs, Cost-efficiency and Cost-effectiveness

Costs are often restricted to implementation or management costs. For this reason the taxation instrument is often a cost-efficient and cost-effective one.

4.4.7 Levels of Evaluation Effort

The simplest evaluation will be one mainly based on reports concerning tax revenues or reimbursements, while a more comprehensive one will also pay attention to more subjective information from the management team and from the organisations and/or persons that are subjected to these taxes.

Table 4.7 Evaluation activities associated with different levels of effort: taxes

Level A **Comprehensive Evaluation**	**Level B** **Targeted Evaluation**	**Level C** **Programme Review**
Development of programme theory and specification of indicators		
Literature review. Background to the legal procedure for the tax system, the specific procedure as well as related systems.	Literature review. Background to the legal procedure for the specific tax system.	Literature review.
Characterisation of programme activity		
Reports on tax revenues and/or reimbursements. Management reports. Programme management interviews.	Reports on tax revenues and/or reimbursements. Management reports. Programme management interviews.	Reports on tax revenues and/or reimbursements.
Estimation of changes in indicators		
Reports on tax revenues and/or reimbursements. Sample surveys on awareness levels and barriers to making use of the tax reductions.	Reports on tax revenues and/or reimbursements.	Reports on tax revenues and/or reimbursements.
Baseline development/estimation of net impacts		
Develop a baseline in the evaluation and include impacts (positive and negative) of other policies.	Develop a simple baseline in the evaluation.	
Estimates of energy savings and emission reductions		
Include usage changes in the calculation.	Include usage changes in the calculation.	Calculation based on key figures
Estimates of costs		
Costs from the perspective of the government and the consumer/investor Use of cost-efficiency and cost-effectiveness indicators.	Costs from the perspective of the government and the consumer/investor.	Costs from the government's perspective.

4.5 Financial Arrangements

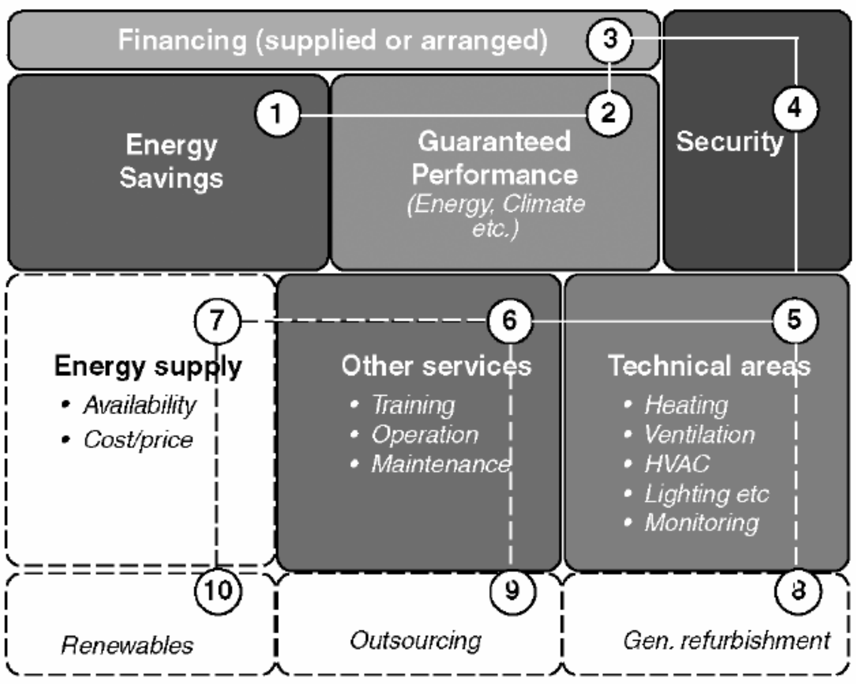
The group 'financial arrangements' will always include actions by additional (market) organisations willing to provide reduced-interest loans, financing guarantees, or a third-party financing facilitation. All contain a similar mechanism: financing of energy efficiency measures using special loans that have to be paid back based on energy savings.

4.5.1 Policy Measure Theory

Policy Measure Domain

Reduced-interest loans, financial guarantees and third-party financing initiatives focus on consumers who invest in energy-efficient techniques and equipment. Loan programmes, for example, help small businesses to finance fixed assets, acquire land, buildings, and machinery, as well as to construct, renovate, expand existing facilities and implement leasehold improvements. These programmes help meet the financing needs of creditworthy small businesses, including manufacturers, that are unable to obtain traditional commercial loans by reducing the risk of traditional lenders.

Figure 4.3 Fundamental elements and most frequent and additional service in energy performance contracting



Source: Westling, 2003 page 18

Guarantee funds for (long-term) energy efficiency investments provide a guarantee for a certain percentage (sometimes up to 70%) of borrowings contracted by small businesses from any bank. Investments requiring long-term borrowing can benefit from such guarantees. Soft loans are those offered at subsidised interest rates, i.e. at rates lower than the market rate. An advantage is that they can be easily implemented by banking institutions. Research indicates that programmes in the residential sector generally have not succeeded because few consumers will take on debt simply to save energy. Yet financial support has promoted business investment in energy-efficient technology and accelerated process innovation. So these measures are mostly targeted towards commercial and non-commercial organisations.

Third-party financing and energy-savings performance contracting often have the same meaning. This arrangement means that a private, semi-public or even a public authority with better financial capabilities will take on the financial risks. They offer the financing in combination with a range of services and product (see Figure 4.3)

Policy Measure Hypothesis

The main objective of the soft loans and/or subsidies is to raise the financial return rates of energy conservation projects to commercially attractive levels. Because pollution prevention and energy efficiency focus on process improvements and waste elimination, they involve capital investments and process changes that are often profitable. Nevertheless the investment is not undertaken due to lack of capital. A financial arrangement overcomes this barrier.

In general a high-risk loan and investment in an energy efficiency measure might be judged risky by a bank, which will have a limited debt proportion, high interest rate and a short repayment time. This is a barrier to investment. A financial guarantee will lower the risk level of the loan and so make it more accessible. By providing incentives, such as loan guarantees, an additional objective can be to attract private lender participation and make more capital available overall.

Reduced interest rate loans or using assistance programmes that reduce loan underwriting and documentation costs, reduce the borrower's costs of financing, making capital more affordable.

Homeowners may wish to improve the energy efficiency of their homes, financing energy improvements through home equity loans, home improvement loans, credit card borrowing, signature loans, or other forms of credit. Reduced utility bills allow a more costly mortgage and so increase the number of homeowners.

4.5.2 Specification of Indicators

Input indicators refer to the monies involved in the financial arrangements.

Examples of **Output** indicators are the:
- number of loans or guarantees given;
- difference in reduced interest and normal interests;
- amount advanced;
- expected energy savings;
- ratio investment/loan.

Examples of **Outcome** indicators include: awareness of energy-efficient products, positive attitudes towards the new energy efficiency products, and changes in future decision-making processes.

The **impact** indicators are (calculated) energy savings and CO_2 reduction relating to the financial arrangements.

Third-party finances, especially with an energy performance contact, generally have a number of specified indicators in the contract (see Table 4.8).

Table 4.8 Important sections in EPC contracts

1. Introduction	Purpose Type of contract 'Shared Savings' or 'Guaranteed Savings'
2. Included RFP (Request for Proposal) and proposal documents	
3. Reference to General Contract Conditions	National, international
4. Energy-saving measures	
5. Financial grade audit	
6. Promise of guarantee	*1. Performance*: - Energy savings: electricity, heating, cooling, etc. - Other criteria: air quality, etc. *2. Operation, maintenance* *3. Investment volume*
7. Options/Other services	Training Outsourcing Energy supply
8. Time schedule	Completion date Length of contract Years after take over Inspection
9. Payments	Different levels and percentage sharing in relation to savings
10. Securities	For implementation, repair, performance, insurance
11. Measurements & verification	Baseline – Adjustments – Weather – Occupancy – How? – By whom?
12. Conditions	Responsibility for pre-audit in case of non-proceeding after detailed engineering
13. Others	Law – Language – Disputes – Cancellation – Force majeure

Source: Westling, 2003 page 33

4.5.3 Development of Baselines

Organisations that provide the loan or finance want to see their money returned. So they will always have some kind of baseline with which to compare the loans: existing energy consumption by fuel, existing energy costs, measures to be installed, expected energy consumption savings from installation, expected cost savings by fuel, expected carbon savings (often based on previous years).

A third-party finance will always include a baseline prior to the finance.

4.5.4 Assessment of Output and Outcome

The financial organisations will often assess the outputs and outcomes of the reduced-interest loans, financial guarantees and third-party financing. They will pay attention to these, as they are important indicators for the return of the loaned or invested money.

4.5.5 Assessment of Energy Savings and Emissions Reductions

Expected savings are generally recorded according to the type of fuel, where total savings are recorded:

- energy-saving GWh/a: calculated by using engineering methods;
- carbon-saving tC/a: calculated by making assumptions on the fuel mix.

Energy savings are crucial, especially in third-part financing, as this determine the return on investment.

4.5.6 Calculation of Costs, Cost-efficiency and Cost-effectiveness

Policy cost effectiveness (€/tC) (based on government funding).
National cost effectiveness (p/kWh) (based on funds from government, partners and customers).

4.5.7 Levels of Evaluation Effort

The government is generally involved in the financial arrangements involved, but the major interests are in the market. This is why a program review will often be conducted (plus sometimes a targeted evaluation). A comprehensive evaluation is seldom made, for the simple reason that if the instrument is not working in a proper way, this will already show up during the implementation.

Table 4.9 Evaluation activities associated with different levels of effort: financial arrangements

Level A **Comprehensive Evaluation**	**Level B** **Targeted Evaluation**	**Level C** **Programme Review**
Development of programme theory and specification of indicators		
	Literature review. Expert interviews.	Literature review.
Characterisation of programme activity		
	Management reports on subsidy and grant programmes. Programme management interviews.	Management reports on loans.
Estimation of changes in indicators		
	Analysis of programme reports.	Analysis of programme reports.
Baseline development/estimation of net impacts		
	Assess the available baseline and adjust as necessary.	Re-use the available baseline in the evaluation.
Estimates of energy savings and emission reductions		
	Calculation based on key figures.	Calculation based on key figures.
Estimates of costs		
	Costs from more than one perspective (e.g. government, finance provider and investor). Use of cost-efficiency indicators.	Cost from one perspective (e.g. energy user).

4.6 Ensuring Minimum Market: Bulk Purchasing and Technology Procurement

Organisations may aggregate large orders of energy-efficient equipment (bulk purchasing) to receive favourable pricing from manufacturers. These price reductions are then passed on to the final customers purchasing the equipment.

Technology procurement may be characterised as an entire acquisition process aimed at directly stimulating innovation. It is not exclusively associated with any particular form of contract, though it is closest to design/build contracting, with functional requirements or performance criteria and functional procurement.

Bulk purchases and technology procurement in combination with other measures, such as labelling and information. Technology procurement in particular is more likely to lead to a permanent change in the market if it is linked, in a planned way, with other concurrent or subsequent actions.

4.6.1 Policy Measure Theory

Technology procurement has been defined as: ‘A process, through which a commodity, service or system is procured, and for which development of new technical solutions is essential in order to meet the requirements of the buyer. The technical development work, being part of the process, may concern the application of advanced technology, but also minor stages of development as well as product modifications. The development work may concern the product, the system or the production process for which it is developed’. (Westling, 2000)

Policy Measure Domain

Despite these common threads, perhaps the most important lesson is that each technology procurement project will have its own unique requirements, resulting from specific barriers and opportunities. (Ten Cate 2002)

Policy Measure Hypotheses

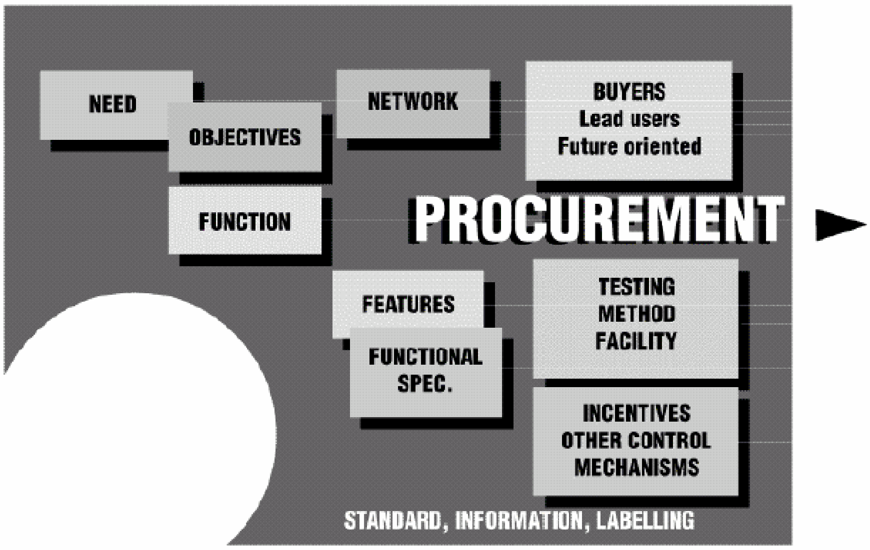
We present several example hypotheses relating to the phases of the procurement process and the principal stakeholders involved. The work carried out the IEA DMS Agreement Task III on technology procurement (see Figure 4.4), differentiates between four phases and four stakeholder groups. These four phases are:

- Preparation.
- Tendering (specification, bid, selecting manufacturer(s)).
- Development (prototype).
- Market acceptance (penetration, target values, labels, market transformation).

and the main stakeholders are:

- Buyers.
- Manufacturers.
- Governmental organisations, energy agencies, leading buyers.
- Supporting organisations/activities (for market acceptance, market transformation).

Figure 4.4: Technology procurement, some important steps



Source: Westling, 2000

The hypotheses for technology procurements are related to the following topics:

- Interaction with buyers and manufacturers: apparently extensive interaction is needed for success.
- Non-energy benefits: these are essential in attracting buyers.
- Participation by manufacturers: specific incentives are important.
- Selection of (and targeting within) the right projects for technology procurement.

The programme development process should be **buyer-driven**. The interests of buyers, their issues and concerns, market perspectives and willingness to buy must largely determine the development and design of a programme. Even though programme development should be buyer-driven, developers should nonetheless seek out suppliers and distributors to learn about their perceptions and motivations and to communicate effectively with existing players, especially companies that might perceive the programme as a competitive threat. In procurement projects aimed at products that have relatively low capital requirements for starting new production, small manufacturers can be important agents of change in the market.

The choice of target technology can strongly influence the success of a procurement project. Target technologies are more promising if they meet the following requirements:

- Products or features are attractive to a large number of motivated buyers.
- Products or features are not already widely available.
- Products are standardised and mass-produced, not custom-designed.
- More than one supplier is in a position to compete for the procurement.
- Desired changes in products or processes are not so fundamental that they require long lead times for R&D.
- The technology advances the developer's strategic goals, e.g. reduced energy consumption.
- The technology brings with it partners and allies to help promote purchases.

4.6.2 Specification of Indicators

Input indicators refer to the monies involved in the programme management and monies used for getting information to the customers and buyers (groups), as well as man-hour capacity. For procurement, an additional input indicator could be the involvement of trusted institutions that are recognised for objectivity, consumer interest, or technical expertise in developing and implementing technology procurement programmes, as these can contribute significantly to their success, although credible independent testing and evaluation of products can be costly.

Output indicators generally refer to the number of sold (awarded) products, if possible specified according to target groups, Additional indicators for technology procurement include:

- Number of companies trying to win the award.
- Number of sold (awarded) products.

Examples of **Outcome** indicators include awareness levels of energy-efficient products, positive attitudes towards the new energy efficiency products, and the proper use of the product, which is important as outcome indicator as it has considerable influence on the value of impact indicators.

The **impact** indicators are (calculated) energy savings and CO_2 reduction related to the subsidised products and measures. As both policies also intend to change the market situation, an outcome indicator could also concern the changes in market shares for the supported products versus the non-supported one. Additionally, indicators could also include the change in % market share or the pooint where the energy-efficient product becomes competitive.

4.6.3 Development of Baselines

The baseline assumption is often quite simple: without the programme the product would be commercially launched into the marketplace much later and slower (technology procurement) or the bulk purchased products are additional to the market. The assumptions made in the energy savings calculations are also used as part of the baselines.

4.6.4 Assessment of Output and Outcome

The IEA DSM INDEEP database and the INDEEP analysis 2004 report contain information on outputs and outcomes for case examples on bulk purchase. Please refer to htttp://dsm.iea.org for more detailed information.

Unless circumstances strongly indicate otherwise, it is preferable to make more than one award in response to competitive applications.

If market and technology conditions allow it, an initial phase application can be very useful in identifying potential suppliers and buyers, appropriateness of specifications, and functionality of programme logistics.

Technology procurement programmes that depend on sales to large volume buyers, particularly government agencies, should be designed to allow a long time period (at least two years) for the target buyers to purchase the product.

4.6.5 Assessment of Energy Savings and Emissions Reductions

The calculated energy savings are based on the number of products sold and average usage time, and are based on product specification versus the selected reference product.

4.6.6 Calculation of Costs, Cost-efficiency and Cost-effectiveness

For bulk purchases, costs are restricted to implementation or management costs, including information costs. Technology procurement projects rely heavily on guaranteed sales or exclusive access to large financial awards (SERP, Apartment-Sized Refrigerators). There are also examples as DOE's clothes washer and sub-CFL programmes that demonstrated that it is not always necessary to attract aggressive bids, especially when attempting modest incremental improvements in technology, not huge leaps forward.

4.6.7 Levels of Evaluation Effort

Table 4.10 Evaluation activities associated with different levels of effort: ensuring minimum markets

Level A **Comprehensive Evaluation**	**Level B** **Targeted Evaluation**	**Level C** **Programme Review**
Development of programme theory and specification of indicators		
Literature review. Expert interviews.	Literature review. Expert interviews.	Literature review.
Characterisation of programme activity		
Analysis of programme records. Programme management interviews. Interviews with applicants (technology procurement).	Management reports. Programme management interviews.	Management reports.
Estimation of changes in indicators		
Analysis management reports. Sample surveys on buyers and/or manufacturers.	Analysis of management reports.	Analysis of management reports.
Baseline development/estimation of net impacts		
Develop a baseline in the evaluation and include free-riders and market segmentation or market changes.	Develop a simple baseline in the evaluation.	Develop a simple baseline in the evaluation.
Estimates of energy cavings and emission reductions		
Include a range in the calculated energy savings, taking into account selected adjustments such as free-riders or rebound. Include market changes into the overall	Include changes in use in the calculation.	Calculation based on key figures.

Level A Comprehensive Evaluation	Level B Targeted Evaluation	Level C Programme Review
estimates on energy savings.		
Estimates of costs		
Costs from different perspectives (government, consumer, producer).	Costs from more than one perspective (e.g. government and consumer).	Programme costs.

4.7 Conclusions

The economic incentives should overcome financial barriers to implementing energy efficiency measures. But in addition to this, when specifying the domain, the main elements shown here for the information policy and measures are all equally important: targeting the programmes towards the selected groups and to the information mechanism.

For the price-reducing policy measures the baselines are seldom included in the programme development phase and, as a result, the outcome and impacts are generally overestimated. For the assessment of the energy savings we refer, for example, to another study within the IEA DSM Agreement, the INDEEP database and the analysis report on this database.

We group several economic incentives together into ‘financial arrangement’ and for this group we conclude that a baseline is available in most programmes, as this is important for forecasting the repayment of the loan, a financial guarantee or the payback period for third-party finance.

Technology procurement (included in the group ensuring a minimum market) has a well-documented system of hypotheses, e.g. as a result of the procurement task under the IEA DSM Agreement. The baseline is also quite simple: the product would be put on the market much later and slower. This should make implementing an evaluation using our seven key elements at a targeted level (B) not too expensive.

5. EVALUATION OF VOLUNTARY AGREEMENTS POLICY MEASURES AND PROGRAMMES

5.1 Introduction

Volume II includes three case examples on evaluated Voluntary Agreements. Results from research projects and workshops on Voluntary Agreements in Europe were also used in preparing this chapter. More detailed information on Voluntary Agreements in Europe and general analyses are generated by three working groups: CAVA (Concerted Action on Voluntary Approaches), NEAPOL (Voluntary Agreements Policy Lessons to be Learned) and VAIE (Voluntary Agreement Implementation and Efficiency).

Table 5.1 Subcategories and case examples for voluntary agreements policy measures and programmes

Subcategories for Voluntary Agreements	Case examples	Country
Industrial companies	Canadian Industry Program for Energy Conservation (CIPEC)	Canada
	Voluntary Agreements	Korea
	Voluntary Agreements on Industrial Energy Conservation 1990 - 2000	Netherlands
	Eco-energy	Sweden
Electricity production, transformation and distribution companies		
Commercial or institutional organisations		

Information from these three case examples is used in various sections to illustrate elements such as output indicators, development of baselines and energy savings.

This chapter starts with a section dealing with the objectives and main types of information policy measures and programmes, but is limited to the strong compliance Voluntary Agreements. For these we present the main evaluation topics, structured by the key analytic elements as discussed in Chapter 1:

5.3 Statement of policy measure theory

5.4 Specification of indicators for evaluation

5.5 Development of baselines for indicators

5.6 Assessment of output and outcome

5.7 Assessment of energy savings and emissions reductions and other relevant impacts

5.8 Calculation of costs, cost-efficiency and cost-effectiveness

5.9 Choice of level of evaluation effort

Based on the case examples and the discussions at the experts meeting, the authors feel that the following issues need special attention in the evaluation of voluntary agreements policy measures and programmes:

- The baseline is a crucial element in the evaluation and should therefore not only be developed for situations without the voluntary agreement, but should also be the main choice for the energy efficiency or energy intensity baseline (economics of physical values used).
- An additional baseline is recommended, which deals solely with the outcome of the voluntary agreement (i.e. without any other measures being combined with the voluntary agreement).
- Audits are often used for target setting in the first phase of a voluntary agreement. Chapter 3 includes key elements for evaluating audits. During the evaluation design phase, when deciding the level of effort for the evaluation, we recommend discussions on whether the audits should be evaluated earlier, rather than at the end of the voluntary agreement.
- This chapter concentrates on voluntary agreements with several sectors. The evaluation should be modified for the new (second) generation of voluntary agreements, which contain unilateral agreements within a sector or group, or voluntary agreements combined with other policy measures.

5.2 Objectives and Main Types of Voluntary Agreements

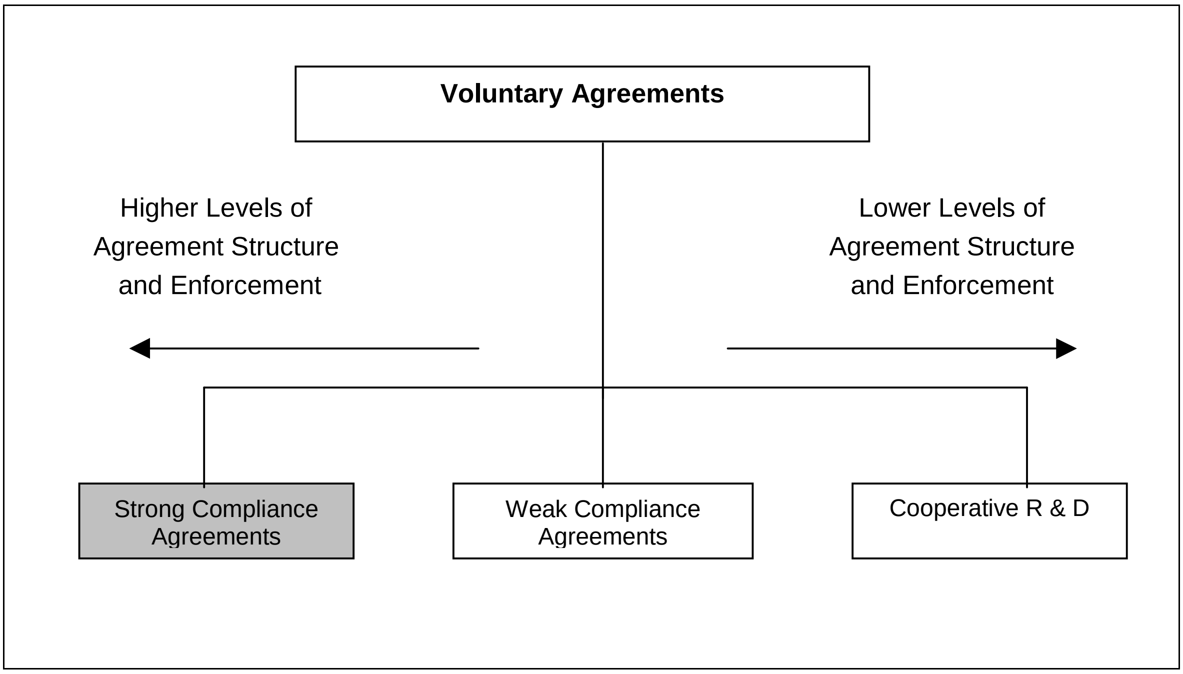
Voluntary Agreements are multi-party programmes aimed at reducing energy consumption among a group of commercial and/or industrial energy users, generally defined according to their industry branch or trade sector. Voluntary Agreements typically consist of the following elements.

- ***Negotiated energy use reduction targets.*** An agency of the national government enters into negotiations with individual energy customers or their industrial association to establish energy efficiency goals. These goals may be expressed in a number of ways: a percentage reduction in energy use per unit of production or facility area from a baseline level; attainment of industry-specific benchmarks for energy use; reduced total energy consumption. The agreement also defines the period over which these energy reduction goals are to be achieved. In some cases, the participants' commitments to Voluntary Agreements are legally binding.
- ***Measurement scheme.*** Most voluntary agreements contain a measurement scheme that specifies the criteria to be used in setting targets, methods for quantifying energy use, and definitions of baselines or benchmarks.
- ***Technical assistance services.*** Many voluntary agreements contain mechanisms for providing technical assistance to the end-users that join the agreement. These services may include technical information dissemination, training, facility auditing and consulting, project design review, and assistance in implementing the measurement scheme.
- ***Sanctions for non-compliance/incentives for compliance.*** Sponsors of Voluntary Agreements have used a wide variety of mechanisms to encourage participating enterprises to meet their obligations. These mechanisms have included:
    - *Threat of regulation.* Some governments have implied that imposition of strict emissions limits or prescriptive equipment standards will result from persistent non-compliance or refusal to participate in Voluntary Agreements.
    - *Withholding of operating permits.* In some countries, compliance with Voluntary Agreements is required for the renewal of environmental permits.

- *Exemption from taxes.* A number of governments, including those in the United Kingdom, Denmark, and the Netherlands, exempt facility owners in selected industries from fuel taxes if they participate in and meet the obligations of industry-wide Voluntary Agreements.
- *Emission trading credits.* A number of governments, including the Canadian federal government, have made provision for tradable 'early action credits' for emission reduction activities within the framework of Voluntary Agreements.
- *Economic Incentives.* Participants in Voluntary Agreements may be offered fiscal incentives such as rebates, credits against corporate income taxes, or reduced-interest loans, for implementing energy efficiency measures.

An International Energy Agency review of national policies with regard to climate change identified three main types of Voluntary Agreement programmes, as summarised in Figure 5.1. These are:

Figure 5.1 Varieties of Voluntary Agreements



1. ***Strong compliance agreements.*** These arrangements are characterised by penalties for non-participation for all facilities over a certain size, clearly-defined energy reduction goals set at the facility (as opposed to the branch) level, clearly defined measurement schemes, and sanctions for non-compliance.
2. ***Weak compliance agreements.*** These arrangements are characterised by voluntary participation, goals that may be set at the facility or branch level, measurement schemes that allow for a variety of reporting methods, and incentives for compliance (without penalties for non-compliance).
3. ***Cooperative R&D.*** These agreements call for industry participation in government-sponsored research and development to identify and develop energy-efficient production and facility management technologies. As exemplified by the United States Department of Energy's *Industries of the Future* programme, representatives of government and

industry jointly develop a 'roadmap' for developing energy-efficient technologies to be followed over a 5-10 year period. The signatories then agree to undertake the various investments and actions required in order to implement the roadmap. There are generally no penalties for non-compliance with these agreements and no direct measurement of energy savings.

This *Guidebook* covers only the evaluation of strong compliance agreements. Readers should bear in mind that the Swedish case example on Eco-energy in Volume II is an example of a week compliance agreement.

5.3 Policy Measure Theory

Policy Measure Domain. Government policies developed to support Voluntary Agreements define the programme domain. These policies generally define the range of facilities targeted by industry branch and size. In the case of the first set of Dutch Long-Term Agreements (MJA-1), the policy specified a goal of bringing facilities (which accounted for 90% of total industrial energy consumption) into the voluntary agreements.

When discussing the programme domain, it should also be noted that the nature of Voluntary Agreements requires that several government agencies be involved in the development and implementation of the policy measure. For example, the development of the Dutch Long-Term Agreements involved representatives from the Ministries of Economic Affairs, Agriculture, Housing, Regional Planning and Environment, and Finance, as well as a number of sub-agencies and contractors.[40] This proliferation of parties involved in programme implementation complicates the documentation of programme outputs, and makes attention to detail in this regard particularly important.

The IEA (1997) combined two characteristics of voluntary actions (degree of structure and legal compulsion) with the programme domain, and concluded that most voluntary actions in the residential, commercial and institutional sectors are less structured. The characteristics of the different types of voluntary agreements are changing over time and it seems that a second set of agreements has now come into action, but that these are more a combination of policy instruments (see Chapter 6).

Policy Measure Hypotheses. The basic hypothesis concerning voluntary agreements is as follows.[41]

1. *Reduction of costs and increases in profitability for investments in energy efficiency relative to other investments.* Many of the components of voluntary agreements serve to increase the attractiveness of investments in energy efficiency relative to other investments. First, governments often agree, sometimes implicitly, to refrain from implementing additional energy and environmental regulations that may lead to

[40] Algemene Rekenkamer. 2003. *Effectiviteit energiebesparingsbeleid in de glastuinbouw.*, (Effectiveness of energy-saving policy in the horticultural sector), The Hague.

[41] Blok, Kornelis et al., 2002. *The effectiveness of policy instruments for energy efficiency improvement in firms.* Utrecht: Universiteit Utrecht, Department of Science, Technology, and Society, Section 5.

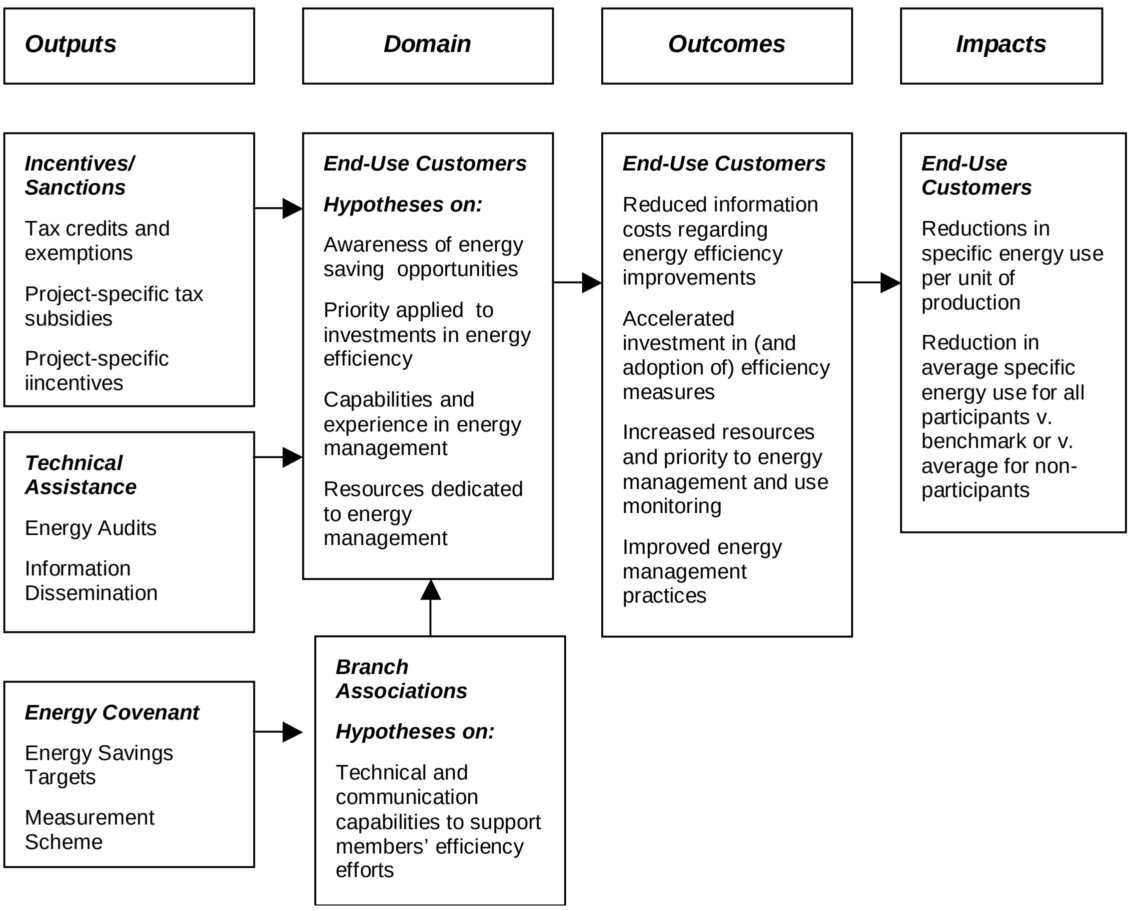
increased costs and risks over the period of the agreement. Second, tax and other fiscal incentives reduce the costs of energy efficiency investments relative to other investments, thus increasing relative profitability. Third, technical assistance and information services provided through the programme reduce information search and learning costs associated with implementing efficiency measures.

2. *Learning.* The relatively long timeframe for participation (5-10 years) and the requirement for measurement and verification of energy usage encourage the adoption and refinement over time of energy management procedures and techniques. In some cases, where the branch (industry) association has taken an active role in programme implementation, technical and management capabilities built up within the association have been used by the participating facilities.

3. *Communication and knowledge diffusion.* Voluntary Agreements can increase investment in energy efficiency through sharing information among group members regarding the efficacy of various procedures and measures. Because this information comes from peers (and competitors), it may have more credibility than similar kinds of information published by government and academic sources.

Figure 5.2 shows the general causal model for these programme effects.

Figure 5.2 Model of Voluntary Agreement Theory



The authors decided to present the Voluntary Agreements as a separate policy measure in this chapter, but it is emphasised that the impact heavily depends on successful embedding in a broader policy mix (Krarup, 2000, OECD 2003, Prince, 2003) and this is included in the next chapter. The specific policy mix depends on the ambition level of the underlying policy strategy and the quality of target setting, which therefore represents the main factors of success or failure. Table 5.2 summarises the supporting policy measures for five Voluntary Agreements. This subject should also be taken into account when developing hypotheses.

Table 5..2 Examples of Voluntary Agreement and Supporting Policy Measures

		Supporting Policies and Measures				
Country	**VA Scheme**	**Government Facilitation of VA Process**	**Audits and Assessments**	**Financial Assistance and Incentives**	**Government and Public Recognition**	**Exemption from Regulation and Taxes**
Australia	Greenhouse Challenge	X			X	
Canada	Canadian Industry Program for Energy Conservation	X			X	
Denmark	Agreements on Industrial Energy Efficiency	X	X	X		X
Netherlands	Long Term Agreements	X	X	X		X
Sweden	EKO-Energi	X	X		X	
UK	Make a Corporate Commitment, Climate Change Agreements	X			X	X

Source: LBNL-52715, Voluntary Agreements, 2003, page 7

5.4 Specification of Indicators

As Table 5.2 suggests, outcome and impact indicators for Voluntary Agreement will focus on:

- Changes in awareness level.
- Changes in knowledge levels.
- Changes in adoption practice levels.
- Changes in energy use.

Table 5.3 shows examples of outcome and impact indicators that have been used for Voluntary agreements. For examples on output indicators please refer to Chapter 3, which contains the output indicators for information programmes.

Table 5.3 Examples of outcome and impact indicators: Voluntary Agreements

Outcome and Impact	**Examples**
Change in awareness level.	• % of targeted enterprises that sign the VA. • % of total sectoral energy consumption accounted for by participants in the VA. • Δ (over time) in percentage of VA compliance plans that mention specific applicable measures.
Change in knowledge level.	• Δ percentage of compliance plans that meet sponsors' technical standards

Outcome and Impact	Examples
	• Δ in energy savings identified in VA compliance plans.
Change in adoption of practices level	• Δ (over time) in percentage of targeted enterprises that participate. • Δ percentage of targeted enterprises that have management staff assigned to the implementation of compliance plans. • Δ percentage of targeted enterprises that implement appropriate measures. • percentage of investments relating to the payback period (> 3 years). • Number, variety, and cost of energy efficiency investments made by participants. • Number and variety of energy management practices undertaken by participants.
Changes in energy use	• Δ (over time) in energy use per unit produced or unit of floor space. • Δ (over time) in energy use per monetary unit of sales • Δ (over time) in average energy use per enterprise in the sector • Total energy savings for documented energy efficiency projects

Assessment of indicators. Evaluators have relied on the following approaches to assess the effects of Voluntary Agreements on participants' decisions with regard to adopting energy efficiency practices and measures.

- ***Participant self-reports.*** Evaluators of the Dutch Long-Term Agreements made extensive use of self-reports from participants to assess the effect (on energy efficiency investments) of the full package of policy measures associated with the agreements. Specifically, participants were asked to rate the effect of the programme on five categories of investments, ranging from 'Good Housekeeping' practices through to purchases of important equipment and CHP systems.
- ***Expert Opinion.*** Some studies have supplemented participant reports with observations of industry experts regarding the likely effects of the programme on levels of energy efficiency investment and/or expected baseline levels of investment.
- ***Cross-sectional analysis.*** Some studies have used interviews with non-participants or data drawn from population surveys of the branches in question to characterise baseline levels of investment in energy efficiency.

Evaluation system. Within the European Union the research project NEAPOL (Negotiated Environmental Agreements: Policy Lessons to be Learned) has developed a comprehensive Voluntary Agreement evaluation system. This system includes four dimensions (see Table 5.4) relevant to measuring the performance or successfulness of a negotiated agreement, which could also be useful for evaluating an energy-based Voluntary Agreement. These four dimensions are:

- Feasibility.
- Capability.
- Impacts.
- Resource development.

Table 5.4 Four dimensions to evaluating environmental agreements

1. **Feasibility**: addresses the question whether the negotiation process resulted in the signing of the agreement
2. **Capability**: addresses the consistency (or match) of the agreement with the underlying policy objectives (the 'specification' aspect) and the extent to which it reinforces, or erodes, these objectives (the 'application' aspect). The specification includes three subdimensions:
   a. Environmental performance
   b. Learning
   c. Economic Efficiency
3. **Impacts**: specified for:
   a. Environmental impact
   b. Cost-efficiency
   c. Competition
4. **Resource Development** addresses the improvements in the policy resource base resulting from negotiating and implementing the agreement.

Source: De Clercq, 2001

For these four dimensions NEAPOL presents its criteria as a questionnaire (that includes a number of statements), a scoring guide (arguments to the applicable scores 1, 3 and 5) and remarks to the scores. Table 5.6 contains an example related to learning. This illustrates the usefulness of this system as a tool for developing indicators or to (re)use the questionnaire for evaluating voluntary agreements.

Table 5.5 Example of measuring the extent to which learning is incorporated into a Voluntary Agreement

I.5	**The agreement contains a clear objective (or objectives) with respect to learning.**
Scoring guide	*5 The agreement contains an explicit objective with respect to learning.* *3 The agreement contains an implicit objective with respect to learning.* *1 The agreement contains no objectives or provisions with respect to learning.*
Remarks	a) Learning could relate to the reduction of information asymmetries (e.g. the dissemination of current best practice, or the collation of existing information and knowledge), or the reduction of shared uncertainties (e.g. the identification of new technical and managerial solutions, or the generation of new information). b) If an agreement does not contain an explicit learning objective, but includes provisions for the collection of new information, sharing of existing knowledge, etc, then this should be interpreted as an implicit learning objective, and would score 2.

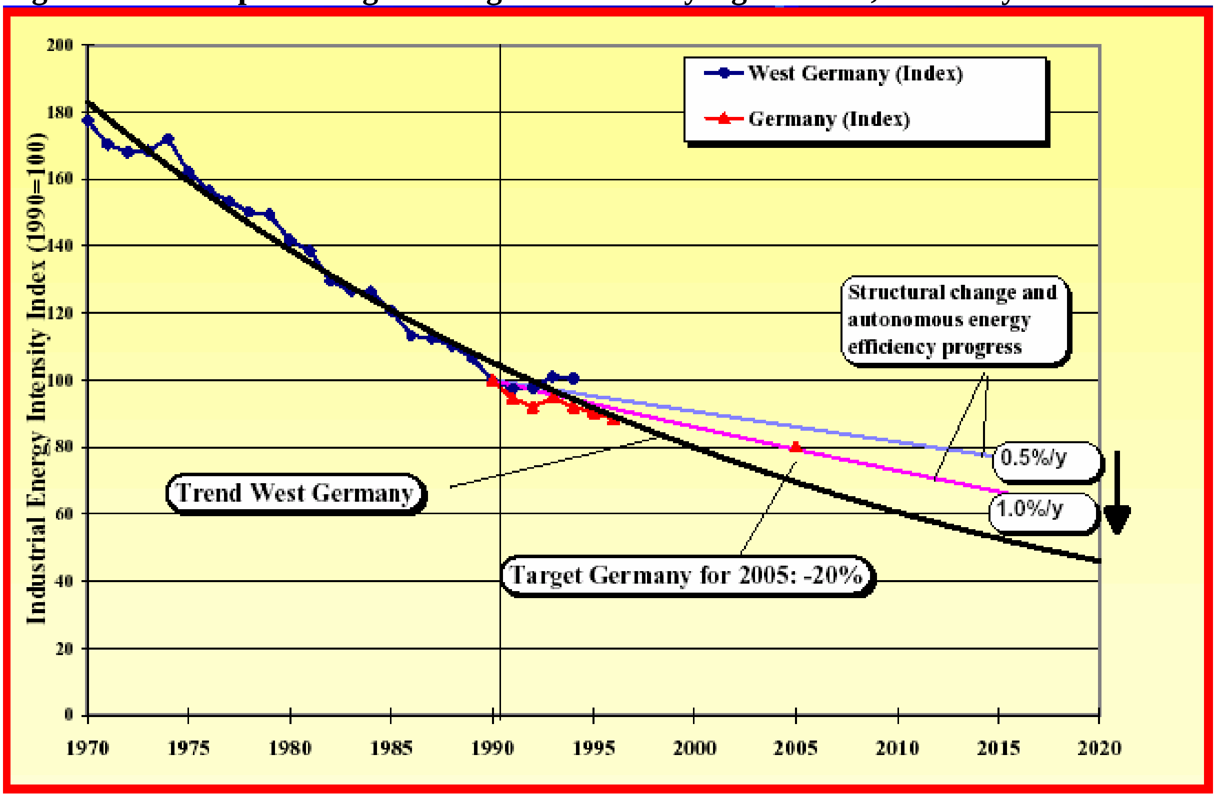
Source: De Clercq, 2001 page 17

5.5 Development of Baselines

For industrial Voluntary Agreements there should be two scenarios: the potential energy efficiency improvement in the period relevant for the agreement for both a 'business-as-usual' and 'with Voluntary Agreement' situation. Using information developed through the assessment of enterprise energy-efficiency improvement potential, as well as information on historical and planned energy intensity, achievable yet challenging targets should be set at the beginning of a Voluntary Agreement.

Baselines for industrial Voluntary Agreements concern energy efficiency or energy intensity. Figure 5.3 contains a baseline example for Germany. The figure presents the historical trend in industrial energy intensity for Western Germany and, in recent years, for a unified Germany. The lines at the top represent the baselines, taking into account structural change (top line) and autonomous energy efficiency progress (middle line). The Voluntary Agreement is targeted towards implementing the future development of Target Germany.

Figure 5.3 Example of target setting for Voluntary Agreements, Germany



Source, Eichhammer, 2003

5.6 Assessment of Output and Outcome

Verification of Outputs. Due to their emphasis on negotiation and monitoring of commitments, Voluntary Agreements generate a significant volume of documentation concerning programme outputs. Key documents include:

- ***Initial Studies.*** Under the approach taken to develop the Dutch Long-Term Agreements (LTAs), the first step in the process, once the government and end-users declare their intent to proceed, is to undertake a study of energy efficiency opportunities in the targeted facilities. This study produces an estimate of the number of facilities in the group, their energy use characteristics, the prevalence of various energy efficiency opportunities, and an estimate of the range of energy savings available in typical facilities. This document serves as the basis for negotiating energy savings targets and other elements of the LTA.
- ***Energy Agreements.*** Energy Agreements represent the basic 'contract' between the government agency sponsoring the agreement (the sponsor) and the participants. The

Energy Covenant usually addresses the following issues: defining energy use criteria (per unit of production, square foot of occupied space, or other denominator), defining a reference year or other benchmark, defining energy-reduction goals versus the reference year or other benchmark. In some cases Energy Covenants are negotiated between the sponsor and an organisation that represents a large group of end-users, such as a trade association or a provincial government. In others, the Covenants are negotiated directly with participating facilities.

- ***Measurement Schemes.*** Measurement schemes outline the methods to be used to measure baseline and future energy usage. They also contain the rules for developing the Specific Energy Consumption (energy use per unit of product or floor space). These rules include conventions for treating the addition or deletion of products from factory runs, adjustments for variations in production and occupancy levels, adjustments for variation in weather, and retirement and addition of floor space, among other issues.

- ***Energy Action Plans.*** Participating facilities develop Energy Action Plans to guide their work towards the energy savings targets. They contain characterisation of energy consumption during the reference year, other energy consumption benchmarks, identification of applicable energy efficiency measures and practices, and a schedule for implementation of cost-effective improvements. Under the Dutch LTA approach, the participants' Energy Action Plans were reviewed by Novem, the government energy and environmental agency charged with oversight of the technical aspects of the programme. If and Energy Action Plan failed to meet Novem's criteria, it was returned to the facility with comment for revision and correction.

- ***Progress Reports.*** Depending on the nature of the Voluntary Agreement, progress reports may be prepared by the participants' representative trade organisation or individually by the participating facilities. Generally, these documents contain information on the energy use monitoring methods used, actual energy use, changes in production or occupancy that could influence energy use, the reference energy use adjusted for current production and occupancy conditions, summary energy efficiency measures (see below), and descriptions of efficiency projects and practices undertaken.

- ***Tax and fiscal incentive applications.*** Applications for tax and fiscal incentives to support measures identified in the Energy Action Plans provide further detail on programme outputs and customer activities.

Assessment of Changes in Programme Effects Indicators. Evaluations of Voluntary Agreements have used the following methods to collect and process information on changes in programme effects indicators.

- ***Analysis of progress reports.*** Most evaluations of Voluntary Agreements have relied heavily upon information provided by participating facilities and trade organisations in annual progress reports. Items covered in these reports include aggregate statistics on changes in energy use and measures implemented. In many cases, these reports are supported by summaries of current energy use, adjusted baseline energy use, and measures implemented by facilities for the participating enterprises. Evaluators of the Dutch Long-Term Agreements found that these documents varied considerably in terms of coverage, documentation of energy calculations, and documentation of changes in production. It was often difficult to make an independent judgement of the energy savings claimed by the reporting facilities.[42]

- ***Customer surveys.*** Evaluators have undertaken surveys of both participating and non-participating customers to assess changes in programme effects indicators and to develop support for cross-sectional analyses of programme influence on customer behaviour. These surveys provide considerable detail on customer behaviour and perceptions of programme influence. However, they are subject to self-reporting errors on both counts.
- ***Secondary statistical sources.*** In most advanced industrial countries, government agencies conduct regular surveys of employment, sales, production, and energy use among manufacturers and other industry branches. The results of these surveys can be used to assess the reliability and accuracy of data developed through the primary collection methods discussed above.

5.7 Assessment of Energy Savings and Emissions Reductions

Analysis of Energy Savings. The analysis of energy savings achieved by Voluntary Agreements departs substantially from approaches that are appropriate for other types of programmes in the following important respects:

- In many cases, programme goals are stated in terms of reductions in facility-level energy use versus a base year quantity, as opposed to savings associated with specific measures.
- Programme goals for energy savings or increases in energy efficiency are often formulated at the branch level, as opposed to the facility or enterprise level.
- Many Voluntary Agreements establish multi-year periods for performance and require long-term monitoring, versus year-by-year savings analyses conducted for other types of programmes.

These programme features give rise to the following challenges in estimating energy savings, taken from the Dutch case example in Volume II[43]:

- Formulation of energy efficiency indices – adjustments for changes in occupancy and production over time.
- Product versus enterprise-level measurement.
- Adjustments for entry and exit of facilities.
- Development of new products.

- ***Formulation of energy efficiency indices – adjustments for changes in occupancy and production over time.*** Given the nature of goal statements for Voluntary Agreements, the indices used to measure energy consumption need to be able to accommodate and reflect changes in volume and composition of production or floor space occupied over time. To meet these requirements, the sponsors of the Dutch Long-Term Agreements devised an Energy Efficiency Index (EEI) that contains adjustments to reflect the differences in the scale and

[42] Blok et al., 2002. p. 91.

[43] Handbook for energy efficiency monitoring, 1999

composition of production between the current and baseline years. The formula used to characterise the energy efficiency of a participating firm is as follows.

A similar equation can be developed for the sector as a whole, with appropriate weighting of facilities for levels of production or occupied floor space.

$$EEI_j = 100 \times \frac{Ej}{\sum_{i=1}^{n} P_{i,j} \times SEC_{i,bj}}$$

Where:

EEI_j = The Energy Efficiency Index for the subject branch in year j.
$Pi.j$ = The physical production amount of product i in year j.

- ***Product versus enterprise-level measurement.*** Programme goals for energy savings or increases in energy efficiency are often formulated at the branch or trade level, as opposed to the facility or enterprise level. In some cases, products are sufficiently homogeneous across enterprises that the sector EEI can be estimated on the basis of measurements made on a product unit production level. For the Dutch LTAs, this was the case for vegetable and fruit processing, vegetable fat processing and domestic porcelain. EEIs for various service industries, such as health and higher education, were developed on the basis of floor space without differentiation for specific types of service. However, some industries such as chemicals, paper, as well as iron and steel produced between 30 and 200 different products. In these cases, product-based EEIs were developed for individual facilities. The sectoral EEI was simply the sum of actual energy consumption divided by the sum of reference energy consumption for the participating facilities.
- ***Adjustments for entry and exit of facilities.*** If facilities enter the voluntary agreement after its initial year, then reference year energy-use information from the new participants needs to be added in calculating the appropriate base year EEIs. Similarly, if facilities exit the agreement, their reference year consumption data needs to be deleted from the reference year EEIs.
- ***Development of new products.*** A number of approaches can be taken to adjusting EEIs to reflect the introduction of new items into a sector's product mix. For some products, specific energy information is available from other enterprises in the same sector and country. Alternatively, specific energies from the same company in other countries may be used. Where production volumes are large, the energy use associated with the new products can be subtracted from the total for the sector to maintain year-to-year comparability. Where volumes are small, the energy consequences can be ignored. In the Dutch case, the introduction of new products had a very small effect on the EEIs in affected branches.

All energy measurements rely on consistent and careful reporting of energy use and production totals by the participating companies. Evaluators of the Dutch Long-Term Agreements found many problems in this regard – some so serious that they prevented effective assessment of the energy effects of the programme in selected sectors. To address these problems the evaluators recommended various improvements to reporting practices, including early negotiation and specification of the items needed and the forms in which they were to be provided, and the application of sanctions if reporting requirements are not met. Finally, where development of EEIs proved infeasible due to the extreme heterogeneity of participating facilities, evaluators approved the use of project-oriented energy savings measurements to support assessment of compliance with the Agreements. To illustrate this problem, we refer to two other publications:

- An analysis of seven Voluntary Agreement programmes found that the programmes could be credited with around 50% of the observed energy-efficiency improvement or emissions reductions (Dowd et al., 2001).
- An OECD review concluded that, for a few examples only, the voluntary policy approach was deemed to have contributed significantly to the fulfilment of a given target. In most cases, factors other than the given voluntary approach seem to explain the major part of any environmental improvement that has taken place (OECD, 2003).

This all indicates that the impacts of Voluntary Agreements (as an isolated policy measure) are difficult to assess.

5.8 Calculation of Costs, Cost-efficiency and Cost-effectiveness

The administrative costs of a voluntary approach, in particular a negotiated agreement, include the costs of :

- Preparing and negotiating the agreement.
- Implementing the agreement.

The costs of preparing and negotiating an environmental agreement differ considerably from case to case. In many instances the costs can be fairly high, e.g. if many parties are directly involved, if the legal status of the agreement is ambiguous, and/or if detailed technical analyses of potential abatement options need to be carried out. For simpler, perhaps less ambitious, agreements the costs can be significantly lower – but this could be to the detriment of the environmental effectiveness of the agreement. A pre-defined framework for negotiating the agreement – similar to that developed in France – might somewhat reduce the 'establishment costs' of new agreements.

Implementing costs also vary considerably. In most agreements, the costs of operating the approach seem modest, but – as in a case where tax obligations would depend directly on the fulfilment of the conditions of the agreement – the operational costs of Voluntary Agreements *can* also be considerable.

There are a number of items relating to the cost efficiency of Voluntary Agreements that can be taken into account, e.g.:

- Marginal abatement costs.
- (Increased) flexibility for firms to find less expensive abatement possibilities.
- Impact of the agreement on the structure (and level of competition within) an industrial sector.
- Impact on technology improvements

There is *a priori* no reason to assume that voluntary policy approaches would serve to equalise marginal abatement costs, and hence minimise the total costs of reaching a given environmental target. On the contrary, the design of these approaches generally tends to make it likely that marginal abatement costs will (continue to) differ (significantly) between the various polluters.

There is a two-way link between voluntary approaches and market structure: The degree of competition in a certain sector can have an impact on the probability that a given voluntary approach will be adopted – and on the rate of participation in that approach. At the same time, the adoption of a voluntary approach can affect the degree of competition within a given market.

Voluntary approaches often include mechanisms that can promote the diffusion of existing technologies – between firms, from research institutes to commercial companies, from companies to relevant public authorities, etc.

5.9 Levels of Evaluation Effort

For Voluntary Agreements, the level of evaluation effort will largely be determined by the degree to which sponsors wish to independently verify claimed reductions in energy consumption among the targeted end-users. The selection of methods used to attribute observed changes in energy use to the programme (versus other influences) could be affected by a number of factors. For example, the percentage of all enterprises participating in the agreements, the extent of systematic differences between participants and non-participants in facility configuration and energy use among participants, and the availability of energy use and production data from non-participants.

Table 5.6 Evaluation activities associated with different levels of effort: Voluntary Agreement Programmes

Level A **Comprehensive Evaluation**	**Level B** **Targeted Evaluation**	**Level C** **Programme Review**
Development of programme theory and specification of programme effects indicators		
Literature review. Expert interviews. Review of records covering the negotiations of the Energy Covenants.	Literature review. Review of records covering the negotiations of the Energy Covenants.	Review of records covering the negotiations of the Energy Covenants.
Characterisation of programme activity		
Review of programme records, Energy Action Plans, progress reports. Corroboration of programme records with individual participants.	Review of programme records, Energy Action Plans, progress reports.	Review of programme records, progress reports.
Estimation of changes in programme effects indicators		
Review of programme records. Participant surveys. Focus groups of participants.	Review of programme records. Participant surveys.	Review of programme records. Focus groups of participants.
Baseline development/estimation of net programme effects		
Self-reports. Participant surveys covering self-reports of programme effects. *Cross-sectional analysis.* Survey of non-participants to assess implementation of energy efficiency measures and compare to	*Self-reports.* Participant surveys covering self-reports of programme effects. *Cross-sectional analysis.* Comparison of energy trends among participants to national trends captured by government	*Self-reports.* Participant interviews covering self-reports of programme effects - small sample or focus groups. *Cross-sectional analysis.* Comparison of energy trends among participants to national

Level A Comprehensive Evaluation	Level B Targeted Evaluation	Level C Programme Review
participants. Comparison of energy trends among participants to national trends captured by government collection of production and energy statistics. *Expert Opinion: Use* expert opinion to 'backcast' investment levels.	collection of production and energy statistics. *Expert Opinion:* Use expert opinion to 'backcast' investment levels.	trends captured by government collection of production and energy statistics. *Expert Opinion:* Use expert opinion to 'backcast' investment levels.
Estimates of energy savings		
Energy use and production/occupancy tracking. Quality control of progress reports at the facility level and branch level. Engineering calculations supplemented by on-site observations if 'project' approach used.	Energy use and production/occupancy tracking. Quality control of progress reports at the facility level (small sample) and branch level.	Energy use and production/occupancy tracking. Quality control of progress reports at the branch level.

5.10 Conclusions

Although this *Guidebook* concentrates on the evaluation of strong compliance agreements, the authors used experience from a wide range of Voluntary Agreements, as well as agreements not dealing with energy but with other environmental issues. In our opinion, the three basic hypotheses are:

1. Cost reductions and increases in profitability for investments in energy efficiency relative to other investments.
2. Learning over a relatively long timeframe.
3. Communication and knowledge diffusion.

The output indicators are similar to most policy measures dealing with the number of agreements, number of participants, number of audits etc. The outcome indicators are also (in general terms) in line with those for communication policy measures, i.e. changes in awareness levels, knowledge levels, plus adoption and practice.

Most attention is given to the impact indicator: the change in the energy efficiency or the energy intensity. In order to assess this indicator, a well-developed and documented method and monitoring system is almost a pre-condition for an evaluation. However, additional information in progress reports, audits or energy action plans is also needed in order to understand this data.

A programme review (level C effort) is seldom conducted when evaluating Voluntary Agreements. Given the long time periods of voluntary agreements and the complexity of the evaluation (also as a voluntary agreement contains elements from other policy measures such as subsidies, audits and communication) a more comprehensive evaluation (level A) will often be commissioned.

6. EVALUATION OF COMBINATIONS OF POLICY MEASURES AND PROGRAMMES

6.1 Introduction

Many policy measures are not implemented as isolated measures, but as part of a mix or package of measures in order to increase the desired effect. This *Guidebook* identifies five main categories of policy measures: Regulation, Information, Economic Incentives, and Voluntary Agreements (VA). Information will generally always be used as part of a package because people need to be aware of what is expected from them and have the proper knowledge before they can act accordingly. Economic Incentives are easily mixed with any of the other measures, while Regulation and Voluntary Agreements tend to exclude each other (although the threat of regulation to back up agreements can be very effective). Table 6.1 lists some examples of commonly applied combinations.

Table 6.1 Possible packages and examples of commonly applied combinations

Measure	*Regulation*	*Information*	*Economic Incentives*	*Voluntary Agreements*
Regulation	Building codes and MEPS for building equipment	Codes and compliance-calculation tools	Building codes and subsidies	X
Information	MEPS and labelling	Labelling, campaigns, and retailer training	Labelling and subsidies	Voluntary MEPS and labelling
Economic Incentives	MEPS and subsidies	Energy audits and subsidies	Taxes and subsidies	Technology procurement and subsidies
Voluntary Agreements	X	Industrial agreements and energy audits	Industrial agreements and tax exemptions	X

This chapter addresses three types of policy measure packages:

4. Regulation, information, and economic incentives.
5. Voluntary Agreements, information, and economic incentives.
6. Market Transformation: economic incentives (technology procurement), information and voluntary agreements.

Volume II includes information on case examples from the participating countries. Case examples from Belgium and Sweden are categorised in the group ‘combinations’. Also other case examples include combinations as outlined in Table 6.2.

Table 6.2 Evaluation case examples containing combinations of policy measures

Policy Measure Package	Case examples	Country
Regulation, information, and economic incentives	MEPS, labelling and rebate programmes MEPS and labelling	Netherlands Sweden
Voluntary agreements, information, and economic incentives	Industrial long-term agreements , energy audits, and tax exemptions Voluntary MEPS and labelling	Denmark, Netherlands Korea
Market transformation	Market transformation	Sweden Belgium

The format used to discuss the evaluation of these combinations will differ slightly from the one used in the previous chapters (using the seven key analytic elements), as this would result in too much repetition. Instead, we focus here on the issues that are specific for each particular combination. The format for evaluating combinations includes:

- Combination theory.
- Choice of indicators and baselines.
- Impact analysis.
- Evaluation effort.
- Critical elements.

Each of these issues will be discussed in separate sections below. The section on market transformation is more comprehensive as this is a special combination of policy measures, while the other two contain policy measures that are discussed in previous chapters.

6.2 Objectives and Main Types of Combinations of Policy Measures

Combinations of policy measures are used to increase the desired effect of such measures, as the effect of a package can – through synergies – be greater than the sum of effects from using each measure separately. In fact, combinations of measures are often applied to provoke a complete transformation of a particular market. This chapter addresses a specific example for each of the two types of packages that are commonly applied today (see also Figure 6.1):

1. Regulation, Information and Economic Incentives

    ↳ Performance standards combined with labelling and rebates

2. Voluntary Agreements, Information, and Economic Incentives

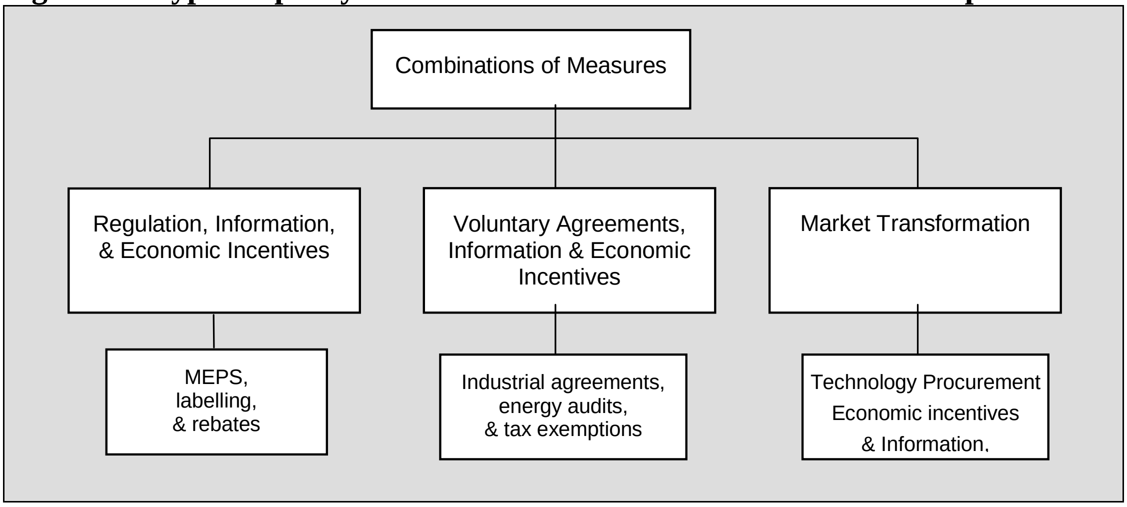
    ↳ Industrial agreements combined with energy audits and tax exemptions

3. Market Transformation Voluntary Agreements, Information, and Economic Incentives

    ↳ Technology procurement combined with information and subsidies (and voluntary agreements

The definitions of each separate policy measure can be found in the previous chapters: using the measures in combination does not change their definitions.

Figure 6.1 Types of policy measure combinations addressed in this chapter

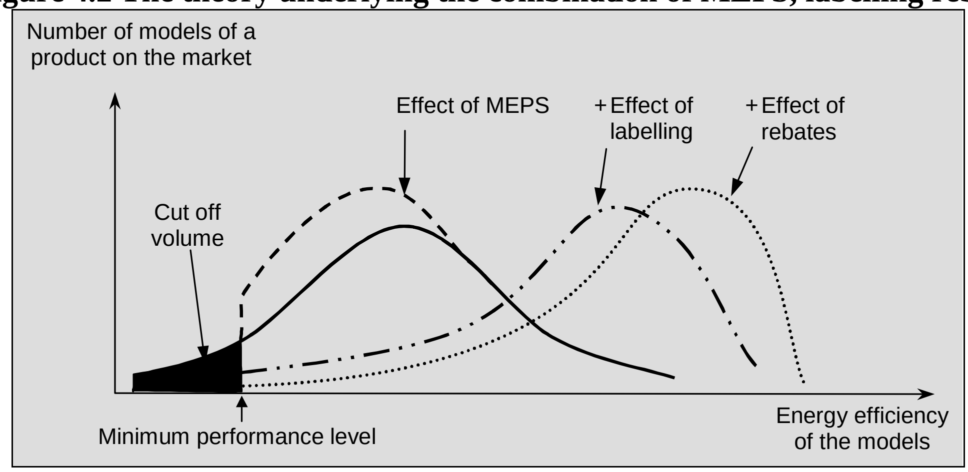


6.3 Combining MEPS, Labelling and Rebates

6.3.1 Combination Theory: How measures can work together

The combination of minimum energy performance standards (MEPS) and labelling is widely used with successful results, and it is generally recognised that MEPS and energy labelling form one of the most cost-effective measures for improving the efficiency of energy appliances. The standards are necessary to remove inefficient products from the market (the cut-off volume), while labelling stimulates technological innovation, as energy efficiency then becomes a competitive issue between manufacturers. Rebates encourage consumers to buy the most energy-efficient products, thereby reinforcing the effect of market transformation. This process is visualised in Figure 4.2.

Figure 4.2 The theory underlying the combination of MEPS, labelling rebates



Source: CLASP (2004)

Note that the separate measures in the packages can have different domains, and combining the measures implies combining the domains. In addition to the hypotheses stated for the measures separately, the combination assumes that the overall effect of the policy measure package will be a market transformation towards energy-efficient products.

The package of MEPS, labelling and rebates can easily be extended to reinforce its impact. For example, training retailers, public awareness campaigns, and technology procurement programmes have all proven to contribute to market transformation.

An interesting question is whether labels and/or standards should be voluntary or mandatory. As Table 4.2 shows, there is a tendency towards mandatory standards and labels, but this approach requires that manufacturers adhere to them, which is not as self-evident as it might seem, as the case of Europe in the 1960s and 1970s shows. However, a voluntary programme may not be effective enough without the threat of the programme becoming mandatory, an approach taken by Switzerland. But Japanese manufacturers routinely meet ‘voluntary targets’ even though Japanese regulations make no mention of enforcement or penalties for not meeting these targets. Apparently, the threat of public disclosure of non-compliance is sufficient deterrent for Japanese companies to make voluntary targets effectively mandatory.

Table 4.2 Mandatory and voluntary labels and MEPS for some products in selected countries

Product	Canada		European Union		Korea		Switzerland	
	labels	MEPS	labels	MEPS	labels	MEPS	labels	MEPS
Air Conditioner - Central	v	m			m			
Air Conditioner - Room	m	m			m	m		
Boilers		m		m		m		
Clothes Dryers	m	m	m				v	?
Clothes Washers	m	m	m	v	m	m	v	?
Clothes Integrated Washer-Dryers	m	m	m				v	
Dishwashers	m	m	m	v			v	?
Freezers	m	m	m	m			v	?
Furnaces	v	m						?
Heat Pumps	v	m			m			
Lamps		m	m		m	m	v	
Radio				v		m		
Ranges/Ovens	m	m						?
Refrigerators	m	m	m	m	m	m	v	?
Space/ Water Heaters- Gas Central	v							
Space Heaters	v	m			m			
Television				v		m		?

Note: List of products is incomplete and only includes comparative labels (not endorsement labels). Source: NAEEC (2001)

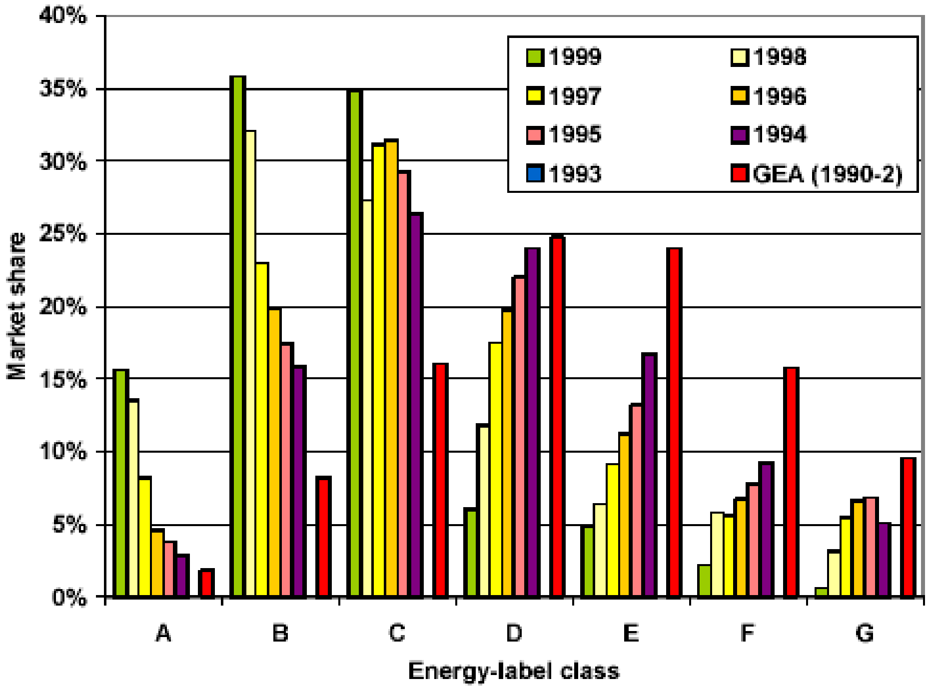
6.3.2 Choice of Indicators and Baselines

When combining policy measures, the choice of indicators and the ex-ante development of a baseline become increasingly important in order to be able to evaluate the package at a later stage. To determine whether a package has led to a market transformation, the right indicators need to be selected. Notably, these indicators include: market share of models of a particular product, but awareness and knowledge of the package among manufacturers, retailers, importers, and consumers are also important. In order to be able to determine whether the package has the desired impact in terms of a reducing energy consumption, proper indicators for end-use energy consumption also need to be selected. Which indicators are used is closely related to the choice of baseline for the package. As energy consumption is generally expected to continue to grow in the future (e.g. as a result of extra appliances per household) the effect of a market transformation on the energy consumption of end-users will not directly become apparent through the energy bills. This (and other issues) needs to be accounted for in the baseline. For example, baselines can be constructed on the assumption that *without* standards and labels, the efficiency of the product would be frozen at the level prevailing at the start of the programme, while assuming a particular value for the increase in average annual energy consumption *excluding* the energy consumption of the product under investigation.

6.3.3 Impact Assessment

The package of MEPS, labelling and rebates aims to affect all stakeholders involved in the life cycle of appliances: designers/producers, importers, retailers and distributors, and consumers. However, the exact impacts can be very difficult to determine accurately, especially when programmes have been in place for some time. The key issue is the baseline, against which evaluators can compare the programme impact. For market shares, this can be done relatively easily if the right data is tracked (markets shares of models that belong to a particular efficiency category). Figure 6.3 shows an example of the market shift towards more efficient refrigerator models in the EU between 1992-1999, as a result of MEPS and labelling. The figure clearly shows a shift toward more efficient models (A, B, and C), while the G-models have almost disappeared from the market.

Figure 6.3 Impact of energy labelling and MEPS for European refrigerators 1992-1999



Source: DEDE 2000

The EU standards and labels for refrigerators resulted in an overall efficiency improvement of 27% between the period 1992-1999. Part of the success of this programme is attributed to additional measures by individual countries, such as rebates, information campaigns, and retailer training.

6.3.4 Evaluation Efforts

The complexity of evaluating a combination of policy measures practically prohibits the use of a simple (level C) approach. Special attention needs to be paid to time lags between the implementation of the packages and the actual impact on society. Mandatory MEPS will induce direct effects (the cut-off volume disappears from the market), but the awareness and understanding of labels among retailers and consumers usually takes considerable time. Intermediate evaluations need to be performed to monitor the progress of key indicators and – if necessary – to redirect programmes.

6.3.5 Critical Elements

As the separate measures have different domains, the hypothesis has to cover these domains and the various time periods. The evaluation should be designed to take this into consideration and the evaluation plan should specify the selected indicators. This selection can result in changing combinations of indicators for the time periods.

As the outcomes and impacts of the combination of policy measures span a period of several years (often over five years) the evaluation should be planned at least at the start of the implementation.

6.4 Combining Industrial Agreements, Energy Audits and Tax Exemptions

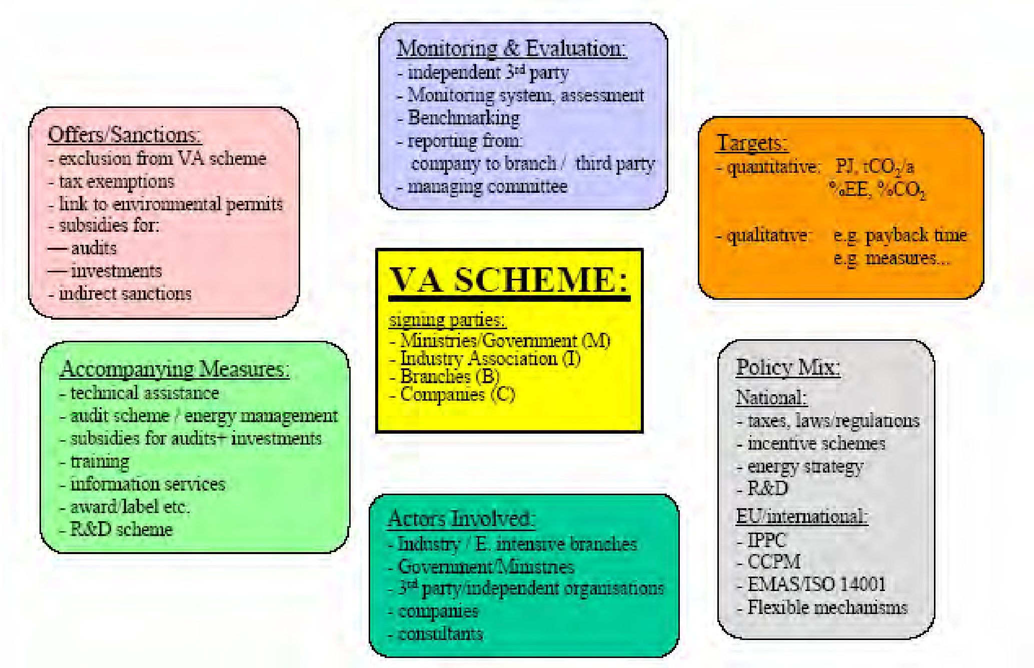
6.4.1 Combination Theory: How measures can work together

The combination of industrial agreements, energy audits, and tax exemptions is used to reduce the energy consumption of the industry and induce innovative energy-efficient production technologies without weakening the competitiveness of the companies involved. The energy audits must provide insight into the measures that offer the highest potential in efficiency gains (at the lowest costs). Tax exemptions will stimulate the companies to actually implement the measures brought forward in the energy audits.

Voluntary Agreements need to be integrated into the existing national policy mix and linked to effective accompanying measures. Figure 6.4 shows the main elements that are considered essential for the successful implementation of a Voluntary Agreement scheme.

In general, existing taxes, laws, regulations, subsidy schemes etc. should be taken into account when designing a Voluntary Agreement scheme. As can be seen from the experiences in several countries (e.g. Denmark, Finland, UK and the Netherlands), taxes are particularly interesting in relation to Voluntary Agreements. They represent a classic 'rod' with the alternative of tax exemptions if entering into a Voluntary Agreement, as allowed under the EU directive on energy taxes. Countries with existing energy audit schemes often use Voluntary Agreements to design a visible and more flexible and effective framework with which to achieve environmental targets. This seems to be the case in Finland and UK, but Denmark also fits in here.

Figure 6.4 Crucial elements in a Voluntary Agreement



Source: Starzer, 2001

A general lesson learned from the analysis of existing LTA (long-term agreement) schemes (Starzer 2001) is that no single element dominates, but that there are several possibilities as to how these elements can be combined to form an effective package, taking into account the existing framework conditions of the respective countries. Of course in reality these approaches often appear in combination. While the tax approach offers a clear offer/sanction mechanism, the audit approach integrates effective accompanying measures to support companies in achieving their targets.

6.4.2 Choice of Indicators and baselines

As previously noted, the choice of indicators and the ex-ante development of a baseline become increasingly important when combining measures, in order to evaluate the package at a later stage. As presented in the chapters ahead, the baseline is a business-as-usual scenario. A large number of evaluations on Voluntary Agreements have shown that the impact on energy efficiency of the single policy measure is not far from the baselines and that the isolated effect of the Voluntary Agreement is not an unbiased process. The combination of policy measures should result in a more visible impact. But a recent Dutch evaluation study of the second generation of industrial Voluntary Agreements concluded that the additional impacts on improved energy efficiency (directly related to the agreement) could only be estimated. Although the companies in the Voluntary Agreement performed better than the average improvement in energy efficiency of the sector, there are several underlying mechanisms that made it impossible to answer the question on the impacts of the policy instrument without doubt (Arentsen 2004, page 18 onwards).

The combination of policy measures should increase the impact. The indicators for the single policy measures are still valid, but when selecting the indicators evaluators should concentrate on those that are most relevant to the combination. A tool in this selection might be the table that Krarup and Ramesohl developed as systematic and comprehensive coverage of possible energy efficiency measures. Particularly with respect to the minor effects in the Voluntary Agreement, the combination should increase these, and indicators in those fields are interesting for the evaluation.

Table 6.3 Options to enhance energy efficiency in industry

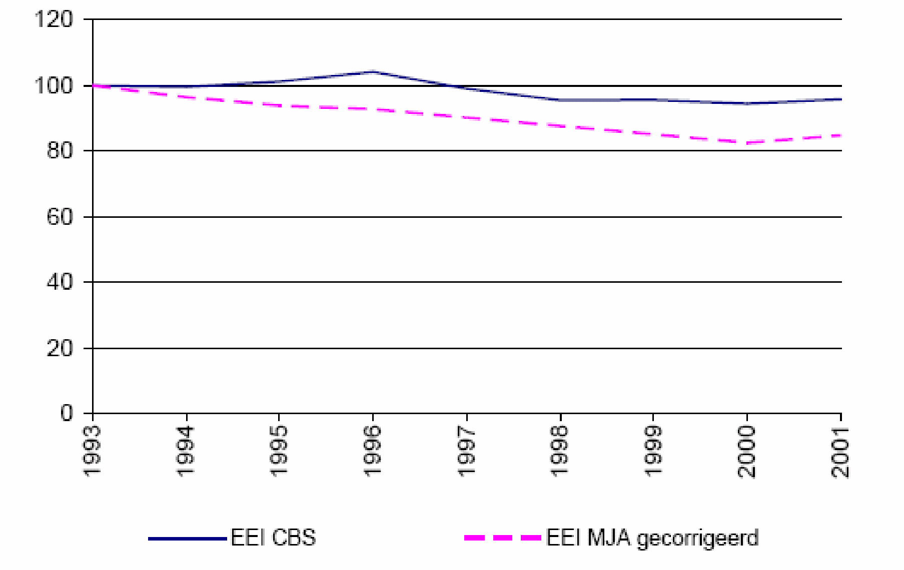
Options	**Requirements**	**Timeframe**	**Impact of agreement schemes**
Change in product design, composition of processed materials and resource use (e.g. thinner and lighter bottles, better recycling etc.).	Strategic commitment and long-term decisions with regard to a change of technical paradigms, process technologies and resource structures.	Long term	Minor effects
Change in energy supply structure (e.g. CHP or renewables).	Strategic commitment and long-term decisions with regard to energy infrastructure and fuel input.	Medium/ long term	Some effects, depending on policy mix (e.g. CHP policy in the Netherlands).
Increased technology innovation.	Strategic commitment and long-term investments into R&D.	Long term	Minor effects
Enhanced investment.	Change in strategic and operative business goals as well as altered decision criteria and procurement procedures.	Short/ medium term	Some effects depending on policy mix (e.g. subsidies) and mandatory requirements (e.g. in Denmark).
Enhanced technology diffusion.	Increased communication, exchange of practical experience, dissemination of best practice and generation of new network links, and even energy-related cooperation of competitors.	Medium term	Some effects, depending on existing cooperation and competition.
Improved energy management.	Integrated approach and systematic search for improvement options, changes in organisational routines, staff empowerment	Medium term	Some effects depending on design of scheme (e.g. integration of audits in Denmark).
Awareness and motivation.	Mobilisation of company stakeholders, provision of information, know-how and expertise, and continuous discussion of the issue.	Short/ medium term	Some effects.

Source: Krarup, 2000 page 39

6.4.3 Impact Assessment

As with the Voluntary Agreements, the impact is often in energy efficiency. Figure 6.5 shows an example taken for a recent Dutch evaluation on Voluntary Agreements. The figure shows the national Energy Efficiency Index (EEI) of one of the industrial sectors (EEI CBS) and the EEI for the companies within this sector that participate in the VA (EEI MJA-corrected). It shows that the companies participating in the agreement perform better than the sector as a whole. As previously indicated, the researchers concluded that it was not yet possible to evaluate the additional impacts from subsidies, sustainable energy or product integration.

Figure 6.5 Example of energy efficiency (1993-2001) for several Dutch companies



Source: Arentsen, 2004 page 23

6.4.4 Evaluation efforts

The combination of policy measures increases the complexity of the evaluation and this requires a higher level of effort for conducting an evaluation. More and more evaluations need to be conducted several times during the multi-year implementation of these policy measures. No detailed evaluation on this combination is known as yet, but as the Dutch evaluation on the new Voluntary Agreement shows (Arentsen, 2004), additional and combined research was necessary in order to evaluate this project.

6.4.5 Critical Elements

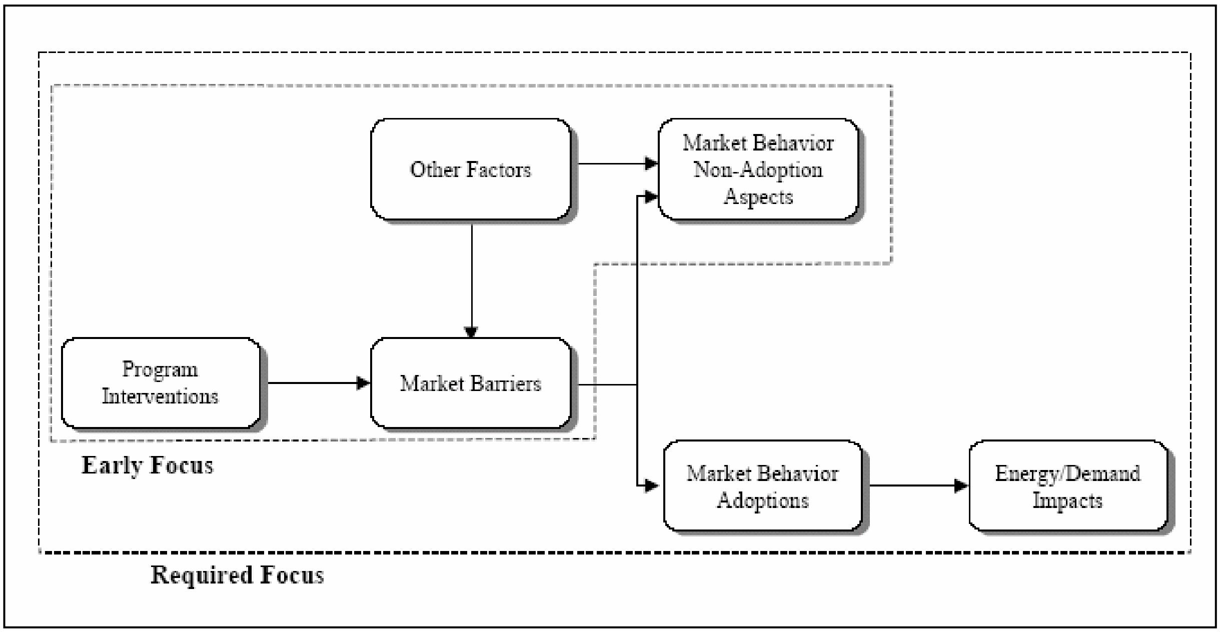
The development of a baseline (in most cases a business-as-usual scenario) concerning the energy efficiency is an important component for the policy programme as well as for the evaluation. As the combination of policy measures should have more impact than the sum of single individual measures (especially the Voluntary Agreement), the evaluation should include more additional information on the explanation of the realised impacts and the underlying assumptions.

6.5 Market Transformation

6.5.1 Combination Theory: How measures can work together

Market transformation is a strategic effort by an organisation to intervene in the market, causing beneficial, lasting changes in the structure or function of the market, and/or practices, leading to increases in the adoption of energy-efficient products, services, and/or practices. These market changes need to be lasting changes, meaning that the changes last beyond any revision or discontinuation of the intervention. In other words, market transformation is a reduction in market barriers resulting from market intervention, as evidenced by a set of market effects, that lasts long after the intervention has been withdrawn, reduced, or changed. The theory (in its simplest form) is illustrated in Figure 6.6. The programme interventions are designed to affect specific identified market barriers. The corresponding changes in barriers induce changes in market behaviour, but not in the total group (early focus). These effects take place over a period of time and involve a variety of market parties and relationships. Changes in adoption lead to energy and demand impacts (required focus).

Figure 6.6 A logical framework of market transformation interventions



Source: Dickerson, 2001 page 6-9

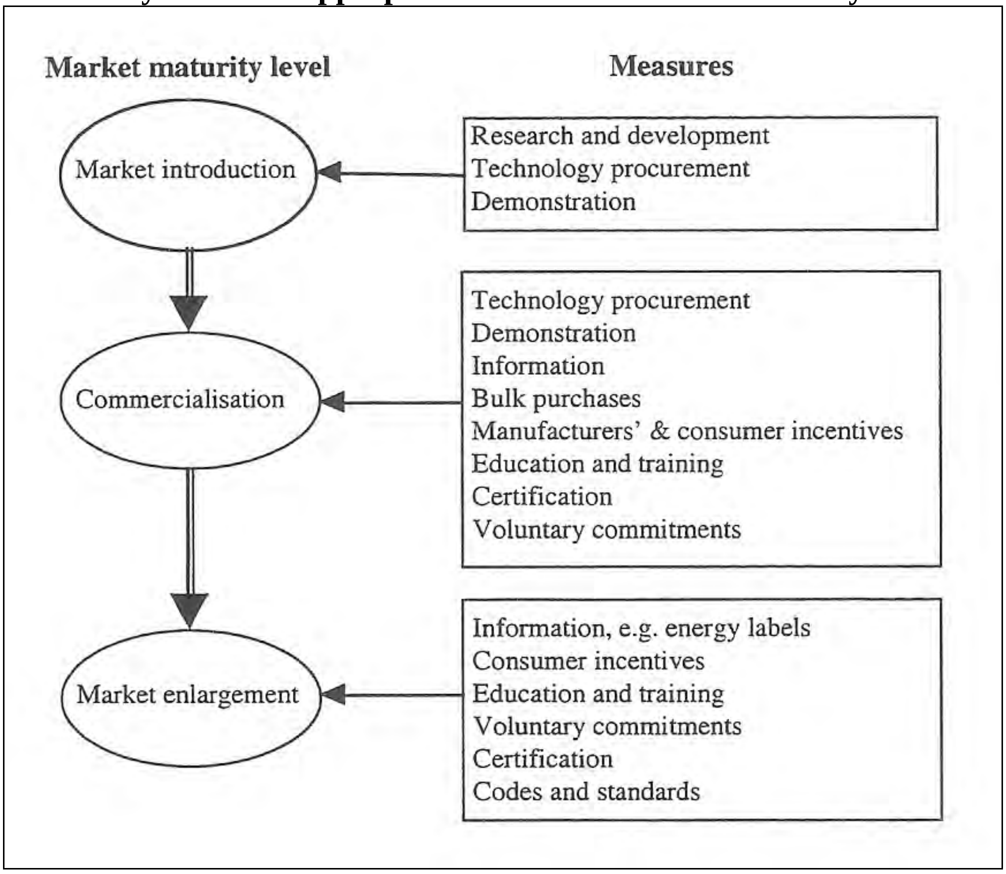
A strategic market transformation intervention will be more likely to succeed if it involves:

- An understanding of the market, market parties, and market channels,
- An understanding of the attributes of the measure or practice and how the measure fits into the market.
- Sufficient energy benefits to warrant the risk and cost of intervention.
- Sufficient private benefits available to interest private-sector collaborators.
- The ability to leverage other stakeholders in the market who, in pursuing their own goals, will help administrators achieve theirs.
- A reasonable logic that ties together these elements into a defensible projection of long-term sustainability of market impact.

This is also the main reason that, for market transformation, there has to be a combination of market intervention instruments such as technology procurement, rebates or subsidies, information, strategic alliances (sometimes formalised in Voluntary Agreements) etc. Figure 6.7 shows an example of policy measures that are more appropriate for a specific level of market maturity and that could be combined into a market transformation policy measure.

The advantages of a market transformation strategy include the fact that even small changes in an entire market can have enormous energy efficiency impacts. A change in a code, standard, or standard operating procedure may seem small, but its effect is magnified over the vast number of transactions influenced by this change. While the immediate cost is not negligible, the ultimate cost to the implementing entity per kWh can be dramatically lower than in a single policy measure e.g. economic incentives. It approaches markets in a more unified and congruent fashion than a policy measure, as that is generally aimed at rate class segments instead of market segments. It leverages not only private-sector investment and innovation, but can run in combination with other regions to create more leveraged changes upstream of the consumer. It can also open up opportunities for continuous efficiency improvements by creating changes in the values of market players and adding new market players.

Figure 6.7 Policy measures appropriate at different market maturity levels



Source: Neij, 1999 page V-4

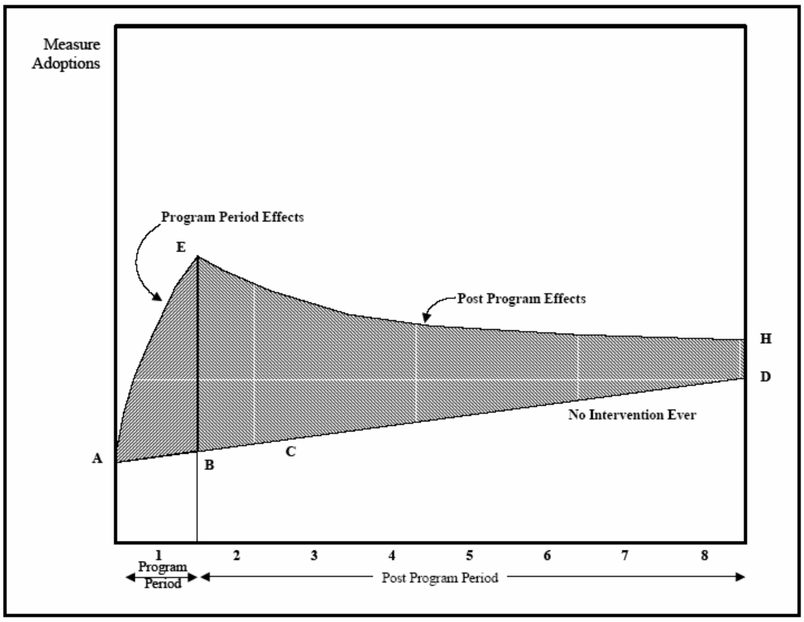
The market transformation intervention is still being developed: a common methodology has not yet been developed (Dickerson, 2001). Also the case example (as included in Volume II) of Belgium (which focused on market expansion rather than technology introduction and commercialisation) and Sweden (market transformation programme to stimulate market penetration) show this clearly. The evaluation of these policy measures is also still being developed.

6.5.2 Choice of Indicators and Baselines

As argued in Chapter 1, impacts are generally defined with reference to a baseline. Because of the nature of market transformation initiatives, the notion of a baseline for market effects is somewhat more complex than for a single policy measure. To accommodate the estimation of market effects, clearer baseline concepts must be developed and the ability to estimate these baselines must be improved.

As market transformation interventions affect behaviour in a lasting way, this change in behaviour will induce adoptions of energy efficiency measures in subsequent periods. Estimating these post-intervention effects requires the ability to forecast behaviour in a dynamic setting. Figure 6.8 illustrates the nature of the problem.

Figure 6.8 Subsequent period effects on adoptions



Source: Dickerson, 2001 page 7-3

Time is shown on the horizontal axis, and adoptions of targeted energy efficiency measures are shown on the vertical axis. For the sake of the illustration, assume that an intervention is operated during a single programme year and then stops. Estimating market effects of multiple year interventions is not a simple task. If programme-year effects are to be estimated, it is necessary to define and measure a baseline for the net changes associated with each year. However, the nature of a baseline is inherently more complex for the market transformation programme than for a singly policy measure to improve energy efficiency. This stems from the fact that market transformation programmes are designed to spawn market effects that last beyond a single programme year. As a result, the baseline for an intervention for each programme year may depend upon the effects of the interventions used in previous programme years. This is one of the main reasons why, in the Belgium case example in Volume II dealing with combinations, no baseline calculation was included in most of the evaluation.

In order to capture the dynamics of market transformation, estimates of market effects must recognise time lags, feedback effects, and a complex set of interactions. Unfortunately, the ability to estimate long-term market effects is, as yet, not particularly well developed. General types of forecasting options that are in use include the Delphi techniques, the analysis of time series market data, and adoption process models. For more information in this field please refer to Dickerson, 2000, Chapter 7, Section 4.

To summarise, the baseline issues in a market transformation policy measure are:

- Time horizon: a longer time horizon for impacts, especially after the policy measure stops.
- Targeted behaviour: affects a wide range of adoption decisions in a particular market.
- Participant focus: a wide range of market parties.
- Overlapping interventions: the baseline for one intervention may need to be adjusted according to the impacts of another.

6.5.3 Impact Assessment

Market transformation evaluation focuses on examining changes in markets at the market level. Given the difficulty of attributing effects from sales data and the associated energy savings estimates, a variety of indicators are crucial evaluation elements for market transformation. These indicators should be a combination of single indicators that provide links to determining whether the programme is leading to (desirable) market changes and whether the interventions have the potential to lead to the anticipated energy savings. In fact, the indicators, outcome and impact, reflect key steps of the programme's logic, and are more important for market transformation than for a single policy measure.

Analysing market transformation impacts is difficult because the:

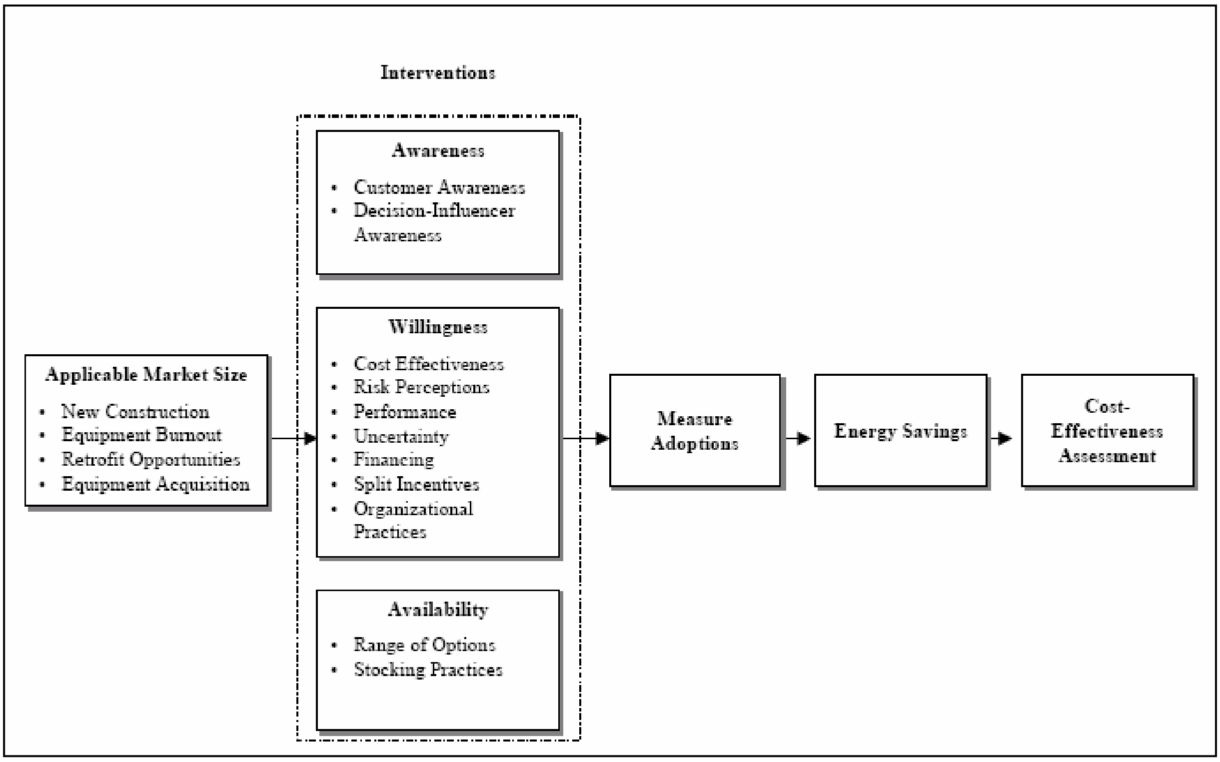
- Participants are often unknown.
- Interventions aspire to cause changes at the market level.
- Interventions tend to be longer lasting, and it takes longer before the ultimate impacts and energy savings are realised.

Nonetheless, the following general model (taken from Dickerson, 2001) illustrates the use of dynamic specifications to generate forecasts of market effects. These forecasted market effects could be used to assess the cost-effectiveness of the policy measures portfolios and specific interventions. Another useful feature of the framework is that it can also be used to simulate the market effects of various intervention options.

The framework characterises the total adoption of an energy efficiency measure as mainly depending on four key factors:

1. **Applicable Market Size**. This reflects the number of opportunities for which the measure is applicable. The characterisation of applicable market size depends on the market event associated with adoptions. Typically, four market events are identified: new construction, replace-on-burnout, retrofit, and appliance acquisition.
2. **Awareness.** In order for an energy efficiency measure to be adopted, decision makers and decision influencers must be aware of the measure and its basic properties. Awareness may be low because of lack of information and/or high information costs.
3. **Willingness.** For decision makers (who are aware of the measure) and are faced with opportunities to install the energy efficiency measure, adoptions depend on their willingness to do so. Willingness depends on a variety of factors, including the number of competing technologies, the cost of the energy efficiency measure, the energy savings obtained through the measure (perhaps differentiated by time-of-use), the energy price, and a host of non-price barriers.
4. **Availability.** The lack of availability of a measure may restrict its adoption by potential adopters who are both aware and willing. Lack of availability may result from manufacturers' practices or problems in the distribution chain.

Figure 6.9 A general conceptual framework to evaluate market transformations



Source: Dickerson, 2001 page 7-11

In the most general terms, the adoption model can be characterised as a chain of these four key factors. This model is in line with the Rogers model on Diffusion of Innovations (Rogers, 1995) and the experience gained within Novem in the Netherlands with a combination of interventions (Egmond, 2003). Also, as the case example for Sweden on market transformation in Volume II illustrates, these elements are often included in more specific evaluations.

More detailed (dynamic) models on intervention effects can be developed and used for each of these four key factors, to calculate separate baselines and impacts. These dynamic frameworks are also useful tools for helping planners and evaluators understand the process by which market interventions (market transformation as well as other interventions) affect selected outcome and impact indicators as well as adoptions of energy efficiency measures.

The development of frameworks offers several distinct advantages:

- The process of constructing this type of model forces planners to make specific assumptions about the impacts of interventions. These assumptions relate to paths of influence of interventions, as well as to the values of the model parameters.
- In general, each model parameter has a specific meaning. Each one provides a means of formalising an assumption or observation about the marketplace.
- Models representing a wide range of phenomena can be developed and tested. For example, both the direct and indirect effects of interventions can be modelled to see how they affect the market. Moreover, if warranted by the context of the analysis, various types of demonstration effects that may enhance sustainability can also be taken into account.
- Once the model is specified in conceptual terms, it can provide a framework for designing evaluations, i.e. it focuses evaluations partly on the estimation of key parameters and selected key indicators.

6.5.4 Evaluation Efforts

The effects of market transformation interventions can be estimated. However, the time horizon of these effects is considerably longer than for most single policy measures. Also, the impacts will be more diffuse and therefore more complicated to trace. Nonetheless, measurement methods can be developed to aid in the tracking and measurement of the impacts, costs and effectiveness of market transformation interventions. These methods have a strong link with the dynamic models that are preferable for the baselines, and this is not an easy task. So the evaluation effort should be at least at a moderate (B) level. For some elements information could be collected while conducting a simple C-level evaluation, especially as input for some detailed models. For the critical elements in the market transformation model (these should be agreed on during the programme development or when developing the baselines) a good evaluation has to a comprehensive one, i.e. the A-level.

6.5.5 Critical Elements

A market transformation intervention is involved in markets over which the programme administrator has limited control: the timing, location, and size of the impact are not controlled. Evaluations should take account of this limitation.

Because it is focused on a dynamic market, it is difficult for an evaluation to track the impacts and process these results into a neat package. Dynamic models form a helpful tool in this field.

Market transformation requires a longer time horizon, and this should be taken into account when planning a multi-year evaluation and monitoring of the impacts.

Market transformation is an intervention that is still under development and the evaluation reflects this situation, so involvement by policy makers during the evaluation process will improve the learning process and the targeting of the evaluation.

6.6 Conclusions

Many policy measures are not implemented as isolated measures, but as part of a mix or package of measures. Combinations of policy measures are used to increase the desired effect of policy measures, as the effect of a package should be larger than the sum of effects of using each measure separately.

Combinations of measures are often applied to provoke a complete transformation of a particular market. The evaluation of market transformation is still developing. It seems that dynamic models could develop as a standard tool for baselines in this field.

When combining policy measures, the choice of indicators and the ex-ante development of a baseline become increasingly important in order to be able to evaluate the package at a later stage. To determine whether a package has led to a market transformation, the right indicators have to be selected and the combination of indicators might change over the time period in which the combined policy measures are implemented.

Evaluation of combined policy measures needs to be more comprehensive and will span a longer time period. So the evaluation efforts need to be at least at a moderate level (B), while (at least for the more critical elements in the combination) a comprehensive evaluation should be conducted to assess the combination. This is why it is particularly important to start the planning of the evaluation in good time, at least when starting to implement the policy measures.

7. CONCLUSIONS AND RECOMMENDATIONS

This last chapter summarises the conclusions and critical points for evaluations as presented in the previous chapters. Section 7.2 presents some conclusions from the case examples as included in Volume II, which deals with the seven key analytic elements and the combination of elements from different types of policies and measures. Section 7.3 includes recommendations for potential follow-up activities.

7.1 General conclusions

When we started to develop this guidebook there were two main objectives:

- To provide guidance for evaluating a broad range of energy efficiency programmes offered by a range of sponsors, such as governments and energy companies, with specific approaches for each major type of programme or policy measure.
- To focus on providing guidance in matching research questions and methodological approaches to programme types and level of ambition.

We hope that these two volumes will help readers to improve evaluations, especially as we have targeted the evaluation process in seven key elements of that the level of ambition is one. Although this volume already contains many examples of theories, indicators and ways used for calculating costs and energy savings, readers might find more detail in the country reports and the case examples (included in Volume II) to help them gain new ideas for their evaluations.

We argue that it is important, even in the development phase of a policy measure or programme, to pay attention to the main components of our seven key analytic elements. A policy measure should already contain the following prior to its implementation:

- A statement of policy measure theory, specified for the domain and effects hypotheses.
- A specification of indicators (output, outcome, impact).
- A baseline.

Based on the country examples, the discussions at the experts meetings and the literature researched, the various chapters contain critical elements and specific conclusions for evaluating the different types and subtypes of policy measures and programmes.

For *regulation* measures and programmes we feel that more attention should be given to baselines. A baseline should be developed prior to the implementation and should include a good description of the assumptions. We also recommend that agreement should be reached on the handling of free-riders and on the expected impact of updating the codes, if this could be foreseen.

Several evaluations (for example see Volume II, the Belgium and Dutch case examples) concluded that the enforcement (and control) of the code by local authorities is a weak link in the chain. Staff should have sufficient knowledge, time, and other resources available to review plans carefully and inspect homes. With the increase of the minimum level the control will become more and more important as the details of the construction (and the correct implementation) often mean the difference between compliance and non-compliance.

Minimum energy performance standards (MEPS) are almost always used in combination with other policy measures, such as labelling and rebate programmes, to improve the desired impact of the measure. If MEPS are used on their own, only the worst performing models are cut off from the market and the mean of the model-performance curve will not change. An evaluation at level C (review) is therefore hard to justify.

The energy savings that could be achieved by using more efficient appliances are often offset by the fact that end-users do not replace their old appliances, but use the new efficient appliances in addition to the older ones, thereby causing an increase in energy consumption instead of a decrease. This rebound effect (replacement versus additional use) should be incorporated as a clear assumption in the models that are used to estimate (baseline) and assess energy savings.

Information policy measures and programmes include a wide range: from general information campaigns and information centres to labelling, education and training, and energy audits. Most hypotheses dealing with general information use some kind of causal model whereby knowledge, attitudes and behaviour are the main features. But in both the policy measure theory and the hypotheses, more attention should be given to the barriers: to what level should these be removed.

There are often no baselines included in the evaluation, especially for information campaigns, while this could easily be implemented.

One should look carefully into the persistence of the information. This persistence is especially important for the relationship between the output and the impact (indicators). In most situations the impacts are specific or almost immeasurable, with a distinctive relation to the policy measures. An exception is the energy audits: here the impact in energy savings is a clear objective and (at least the potential) energy savings are included in evaluations. Following the Sep metric handbook we recommend that the evaluation should pay special attention to the adjustments for realisation rates from audits.

The *economic* incentives should overcome financial barriers to implement energy-efficient measures. However, the specification of the domain should also pay attention to the targeting of the programmes to the selected groups and to the mechanism used in order to inform them.
Elements are often included in the hypothesis, which deal with the assumptions regarding expected market changes and relations with reducing future prices for products. This is one of the main elements in the outcomes of these measures and these assumptions should be researched, not only in the evaluation shortly after the policy has come to an end, but also in the long run.

With the price-reducing policy measures, baselines are seldom included in the programme development phase and, as a result, the outcome and impacts are generally overestimated. For the assessment of energy savings we refer to another study within the IEA DSM Agreement, the INDEEP database and the analysis report on this database.

We have grouped several economic incentives together into ‘financial arrangement’, and for this group we conclude that a baseline is available in most programmes as this is important for forecasting the repayment of the loan, a financial guarantee or the payback period for third-party finance.

When developing the baseline special attention should be given to the question ‘How to handle free-riders and spill-over effects?’ as these may have considerable impact on the conclusion whether (or not) the policy measured had a (significant) impact.

The technology procurement (included in the group ensuring minimum market) has a good documented system of hypotheses, also as a result of the IEA DSM Agreement’s procurement task. The baseline is also quite simple: the product itself would be launched much later and more slowly onto the market. This should make an evaluation using our seven key elements at a targeted level (B) not too expensive.

The baseline is a crucial element in the evaluation of *voluntary agreements* and, for this reason, should not only be developed for situations without a voluntary agreement, but should also be considered the main choice for the energy efficiency or energy intensity baseline (economic or physical values used). We recommend developing an additional baseline that deals only with the outcome of the voluntary agreement and without other measures that are combined with the voluntary agreement.

For outcome indicators, attention should be paid to at least the following three hypotheses:

4. Cost reductions and profitability increases for investments in energy efficiency relative to other investments.
5. Learning over a relatively long timeframe.
6. Communication and knowledge diffusion.

Voluntary agreements generally focus on the impact indicator: the change in the energy efficiency or the energy intensity. In order to assess this indicator a well-developed and documented method and monitoring system is almost a pre-condition for an evaluation. But in order to understand this data, additional information from progress reports, audits or energy action plans is also required.

Evaluations are seldom conducted on a programme review level (effort level C). Given the long time periods of voluntary agreements and the complexity of the evaluation (also because a voluntary agreement contains elements from other policy measures such as subsidies, audits and communication), a more comprehensive evaluation (level A) is often commissioned.

Combinations of policy measures are used to increase the desired effect of these measures, as the effect of a package should be greater than the sum of just using each measure separately. These combinations are often applied to provoke a complete transformation of a particular market. The evaluation of market transformation is still developing. It seems that dynamic models could develop as a standard tool for baselines in this field.

When combining policy measures, the choice of indicators and the ex-ante development of a baseline becomes increasingly important, in order to evaluate the package at a later stage. To determine whether or not a package has led to market transformation, the right indicators need to be selected and the combination of indicators may well change over the complete timeframe that combined policy measures are implemented.

The evaluation of combined policy measures needs to be more comprehensive and will span a longer time period, so the evaluation efforts need to be at least on a moderate level (level B), while a comprehensive evaluation should be conducted to judge the more critical elements of the combination. For this reason it is even more important to start the planning of the evaluation in good time plus at least at the start of the policy measure implementation.

A market transformation intervention concerns markets over which the programme administrator has limited control: the timing, location, and size of the impact are not controlled. This limitation should be included in the evaluation framework. Since it focuses on a dynamic market, it is difficult for the evaluation to track the impacts and bring these results into a neat package. Dynamic models can be helpful tools in this area.

7.2 Conclusions from the case examples

Although the case examples included in Volume II are not representative of all the evaluations in a country, the total number (32) gives an impression of the weak and strong elements in evaluations.

We concluded that the baselines and indicators received the least attention. There are only a few case examples where the baselines received (significant) attention. The low attention rate for the indicators is mainly caused by the fact that, at the start of a policy measure, there were often no *specific* indicators selected and, in particular, the outcome indicators were missing. In almost all evaluations the (expected) energy savings (and related emission reductions) were included, while at the start of a programme these were often not well specified. The energy savings were often based on (estimated) utilisation data and were not adjusted to external parameters.

The baselines often received some attention but there were only a few cases where baselines received more specific attention. This is mainly caused by the fact that the baselines are referred to as the ‘business as usual’ scenario, and this assumption is also included in the calculated energy savings. In almost all the evaluations presented, the output and outcomes are included, but an in-depth analysis is seldom conducted.

The level of evaluation efforts varies throughout the case examples. Four case examples are labelled as depicting a comprehensive evaluation (effort level A), while an additional nine are judged as a moderate effort (level A/B). The level of evaluation effort in seven case examples is indicated as moderate (level B) and two are almost moderate. Nine case examples are considered to be at the lowest level (C), i.e. programme review level.

We also considered which of the evaluated types and subcategories of the policy measures in the case examples contain elements from other categories. The five case examples on ‘regulations’ all contain ‘information’ elements – general information, labelling, audits and education and training – and one also includes demonstration and project subsidy. In total eight elements (other than regulations) are important and/or included in the case examples in the category for regulations. Only four of the information programmes include elements from the ‘economic’ category: three include subsidies, while one contains grants. On the other hand, the majority of the programmes in the case examples combine several information elements. The case examples in the ‘economic’ category almost all include general information as an element in the programme.

7.3 Recommendations

A considerable amount of literature was consulted during the project. Most came from the USA, with additional information from Australia, IEA/OECD, UN etc. and the eight participating countries (Belgium, Canada, Denmark, France, Italy, Republic of Korea, the Netherlands and Sweden). One of the problems we encountered during the project was the fact that the evaluations are in each country's native language. Some evaluations have been translated into English and others contain summaries or conference papers in English. Since all but one of the countries from which experts participated have a native language other than English, most of the references to the case examples are not available in English. A considerable step forward in improving the distribution of the experience gained from evaluations conducted would be to ensure that the evaluation reports all contain an English-language summary.

We also used the expert's knowledge of the national evaluation systems and the handbooks that are currently available or being prepared. In our opinion, these documents contain information that is important for the international community and so they should be translated into English. A good example is the translation of the Dutch Manual for Evaluating Climate Change Policies (2004). However, other important documents e.g. the Danish evaluation guidebook on evaluating energy utility programmes or the Italian guidelines for white certificates, are currently not available in English.

There is a longer evaluation tradition in the USA, where there is also a bi-annual knowledge exchange at a conference dealing solely with evaluating energy efficiency programmes (International Energy Program Evaluation Conference, IEPEC). All over the world there are national and international evaluation communities who are also dealing with evaluation from another angle, i.e. other than energy and energy efficiency. In order to increase this knowledge exchange and the quality level of evaluations, the European Commission or the IEA might consider supporting the development of a specific energy efficiency evaluators network.

The UNFCCC framework countries (and the European Union) must include the impacts of their policies and measures on the greenhouse gas emissions in their National Communications. EU Member States must also report this information to the European Commission. This guidebook could be a step forward in easing the comparison of these impacts. We recommend that the European Commission, DG Environment, and the UNFCCC secretariat research this option in more detail e.g. via a trial to review several national reports using the key analytic elements as introduced in this guidebook.

Almost every handbook on evaluation argues that it is important to discuss the evaluation framework at the start of the implementation and we gladly support this recommendation. At that point the communication process starts and the evaluation also increases in value if the evaluators are involved. They become familiar with the thoughts that policy makers have for selecting a specific measure and their expectations. This will also improve the communication during the evaluation and the output of the evaluation itself.

REFERENCES

Chapter 1: General

An Evaluation of Phare-financed Energy and Environment Programmes, Final Report, September 1999

Australian Greenhouse Office, *Annual report of the National Appliance and Equipment Energy Efficiency Program (NAEEEP) 2003*, March 2004

Bowie, R. and others, *A Swedish handbook for the evaluation of DSM activities*, Proceedings International Conference on Demand-Side Management and Energy Efficiency, Stockholm 1993

California Public Utilities Commission, *Energy Efficiency Policy Manual*, version 2, August 2003,

California standard practice manual: economic analysis of demand-side programs and projects, October 2001

Dutch Ministry of Finance, *Guidance Evaluation research ex post* (in Dutch, titled Handreiking Evaluatie-onderzoek ex post), 2003

Dyhr-Mikkelsen, K. et al., *A Handbook in Evaluation on Energy Saving Activity*, Proceedings ECEEE summer study 2003

European Commission Agriculture Directorate-General, *Guidelines for the evaluation of LEADER+ programmes* (document VI/43503/02-REV.1), January 2002

European Commission, Evaluation Unit H/6, EuropeAid Co-operation Office, *Evaluation in the European Commission, A guide to the evaluation procedures and instructions currently operational in the Commission's external co-operation programmes*, Brussels, March 2001

Hall, N., *Application and Use of the California Evaluation Framework*, presentation ACEEE Market Transformation Conference 2004

Hungary Energy Centre, Energy Efficiency, Environment and Energy Information Agency and Energy Saving Trust (2002) *Inventory of Present Monitoring and Evaluation Systems*, Hungary

IEA/OECD, *An initial view on methodologies for emission baselines: case study on Energy Efficiency, 2000*, prepared by Hagler Bailly Services

IEA/OECD, *Dealing with climate change, Policies and Measures in IEA Member Countries*, 2000

IEA/OECD, *Energy and Climate Change, an IEA Source-Book for Kyoto and beyond*, 1997

IEA/OECD, *Improving Energy Efficiency*, 2002

IEA/OECD, *An initial view on methodologies for emission baselines: case study on Energy Efficiency*, 2000, prepared by Hagler Bailly Services

International Performance Measurement & Verification Protocol, Volume I and II, revised version March 2002, available electronically at http://www.doe.gov/bridge

Jordan, G.B., *Developing Logic Models and Using Them to Define a Balanced Set of Metrics*, handouts workshop August 19, 2003, IEPEC 2003

Laponce, B., Jamet, B., Colombier, M. and Attali, S. *Energy efficiency for a sustainable world*, ICE, Paris, 1997

Ministry of Housing, Spatial Planning and Environment, *Manual for evaluating Climate Change Policies*, Den Haag, 2004

Nilsson, H., *Looking inside the box of market transformation*, Proceedings ACEEE summer study 1996

OECD, *Environment Directorate, environment policy committee, monitoring, reporting and review of national performance under the Kyoto protocol*, env/epoc(99)20/final, June 1999

Skumatz L. A., *Techniques for getting the most from an evaluation: Review of methods and results for attributing progress, non-energy benefits, net to gross, and cost-benefit*, Proceedings ECEEE summer study 2005

SRC International and others, *A European Ex-post evaluation guidebook for DSM and EE service programmes*, 2001

SRI International, AKF and Elkraft system, *Handbook for evaluating energy saving projects* (Handbog i evaluering af energispareaktiviteter) January 2003 (in Danish)

Taylor-Powell, E., *The Logic Model: A program Performance Framework*, University of Wisconsin-Extension paper

Techonopolis, *Evaluation and Impact Analysis of RTDI Programmes*, presentation DG Research, 31 July 2001

Technopolis, *An international review of methods to measure relative effectiveness of technology policy instruments*, final report, July 2001

TecMarket Works and project Team Members, *The California Evaluation Framework*, prepared for the California Public Utilities Commission and the Project Advisory Group, revised version September 7, 2004

USAID, Office of Energy, Environment and Technology, Best Practices Guide: *Economic and Financial Evaluation of Energy Efficiency Projects and Programs*, prepared by Econergy International Corporation Boulder, Colorado

Vine, E. and Sathaye, J., *Guidelines for the Monitoring, Evaluation, Reporting, Verification, and Certification for Energy-Efficiency Projects for Climate Change Mitigation*, LLBNL-41543, March 1999

Violette, D.M., *Evaluation, Verification, and Performance Measurement of Energy Efficiency Programmes*, 1995

Violette, D.M. & Cooney, K., *Retrospective Assessment of the NorthWest Energy Efficiency Allience, Findings and Report*, final report December 8, 2003

Vreuls, H.H.J., *Harmonization of evaluation methods for energy programs within Europe*, Proceedings IEPEC 2001

Vreuls, H.H.J., *How to compare impacts of the same kind of programs to reduce GHG*, Proceedings IEPEC 2003

Chapter 2: Building Codes

Arkansas Energy Office, 1999. *Energy Performance Evaluation of New Homes in Arkansas*. Website: www.1800arkansas.com/energy/files/pdf/Performance.PDF.

Australian Greenhouse Office, 2000. *Impact of minimum performance Requirements for Class 1 Buildings in Victoria*. Common Wealth of Australia. Website: www.greenhouse.gov.au/energyefficiency/buildings

City of Fort Collins, Colorado, 2002. *Evaluation of New Home Energy Efficiency*. Summary Report, City of Fort Collins. Website: www.fcgov.com/utilities/home-study.php

Ecofys, 2004. *Evaluatie van het klimaaatbeleid in de gebouwde omgeving 1995-2002* (Evaluation of Climate Policy in the Built Environment 1995-2002) in Dutch

Jeeninga, H, Uyterlinde, M.A. en Uitzinger, J.,2001, *Energieverbruik van Energiezuinige woningen (energy use in low energy houses)*, ECN-C—01-072, in Dutch

Regional Economic Research (RER), 2001 *Residential New construction Study*, July 26, California

WEC (World Energy Council), 2001. *Energy Efficiency Policies and Indicators – Policy Evaluations*. Website: www.worldenergy.org/wec-geis/publications/reports/eepi/policy evaluation/policy_evaluation.asp

WSU (Washington State Univeristy, Tony Usibelli), 1997. *The Washington State Energy Code: The Role of Evaluation in Washington State's Non-Residential Energy Code*. Website: www.energycodes.gov/implement/ documents/case_monitor.doc

Xenergy, 2001. *Impact Analysis of the Massachusetts 1998 Residential Energy Code Revisions*. Prepared for the Massachusetts Board of Building Regulations and Standards, Boston. Website: www.energycodes.gov/implement/ pdfs/Massachusetts_rpt.pdf

Chapter 2: Performance Standards (MEPS)

CLASP (Collaborative Labelling and Appliance Standards Program. Lead Authors: Stephen Wiel and Jjames E. MacMahon).2001. *Energy-Efficiency Labels and Standards: A Guidebook for Appliances, Equipment, and Lighting*, CLASP, Washington D.C., USA. *This information is also obtainable through the CLASP website:* www.clasponline.org/

DEDE (Department of Alternative Energy Development and Efficiency, Thailand Ministry of Energy, Lead Author: Peter du Pont) 2002. *Policy Options for Improving the Energy Efficiency of Electrical Appliances*

UN ESCAP Ad Hoc Expert Group Meeting on End-Use Energy Efficiency towards Promotion of a Sustainable Energy Future, 18-20 November 2002, Bangkok, Thailand.

IEA/OECD, 2000. *Energy labels and standards*. www.iea.org/textbase/nppdf/free/2000/label2000.pdf.

Koomey, J.G., Cramer, M. Piete, M.A., Eto, J.H., LBL, Energy & environment Division, December 1995 *Efficiency improvements in U.S. Office Equipment: Expected Policy Impacts and Uncertainties*, LBL-37383

Nadel, S, 2003, *Appliance & equipment efficiency stands in the US: Accomplishments, next steps and lessons learned*, proceedings ECEEE summer 2003

WEC (World Energy Council), 2001. *Energy Efficiency Policies Indicators*. Also obtainable from the WEC website:
www.worldenergy.org/wec-geis/publications/reports/eepi/download/download.asp

Chapter 3: Information:

Australian Greenhouse Office, *Evaluation of the Australian Energy Efficiency Standards and Labelling Program, a draft framework*, prepared by Nick Banks, Report No 2002/20, 22-Feb-2002

Australian Greenhouse Office, *Motivating Home Energy Action - A Handbook of what works*, prepared by Michelle Shipworth, April 2000

California standard practice manual: economic analysis of demand-side programs and projects, October 2001

Château, B., *Package of coordinated measures*, Enerdata, presentation ADEME-WEC Workshop on Energy Efficiency Policies,15-16 December 2003,

Chistensen, W. and Espergren, K.A., Audit II, *Topic Report Monitoring and Evaluation*, draft version September 2002

Collaborative labelling and appliance standards program, website section Evaluation appliance standards and labelling programmes
www.clasponline.org/standard-label/development/evaluate/index.php3

Dyhr-Mikkelsen, K. and others, *Evaluation of free-of-charge energy audits*, Proceedings ECEEE summer study 2005

European Commission, *Paper I Defining criteria for evaluating the effectiveness of EU environmental measures,* 1999

Megdal, L., *The draft new Caliornia Evaluation Framework Concerning Training Efforts*, presentation ACEEE Market Transformation Conference 2004

Motiva, *Enhanced and Stimulated Evaluation of Energy Efficiency Projects*, 1999,

Shindler, B.,Cheek, K.A. and G.H. Stankey K.A., *Monitoring and Evaluating Citizen-Agency Interactions: A Framework Developed for Adaptive Management*, USDA, April 1999

Smith, S. and Thorne, J., *An evaluation of the EnergyGuide Label: what we learned,* Proceedings IEPEC 2003

State Energy Programs (SEP*), SEP Metrics Handbook, Evaluation*, August 2001

Taylor-Powell, E., *The logic Model: A program Performance Framework*, presentation at the website University of Wisconsin-Extension Cooperative Extension

Chapter 4: Economic incentives:

Baranzini, Goldenberg, and Speck, *A Future for Carbon Taxes in Ecological Economics*, March 2000

Beer, de J.G. and others, *Onderzoek effectiviteit energiesubsidies* (Research on the effectiveness of energy subsidies), December 2000, in Dutch

Berenschot, *Evaluatieonderzoek Milieu Actie Plan 1991-2000* (Evaluation on the Environmental Action Plan of the energy distribution sector), 2001, in Dutch

Boonekamp, P.G.M., Jeeninga, H. Heinink H., *Effectiviteit energiepremies, analyse voor het huishoudelijk verbruik tot 2010* (Effectiviness of the energy scheme, analysis on the energy use in households till 2010), 2000, ECN-C-00-062, in Dutch

Buddenberg, A and Visser, R., *Strategies for Technologies Procurement in Creating Energy-Efficient Lighting Installations,* Proceedings Right Light 4, Volume 1, 1997

Capros, P. and others, *Climate Technology Strategies 2, The Macro-Economic Cost and Benefit of Reducing Greenhouse Gas Emissions in the European Union*, ZWE Economic Studies 4, 1999

Cate ten A., Harri J., Shugars J. and Westling H., *Technology Procurement as a Market Transformation Tool*, 2002

CEA, *Evaluatie drie tenders EZ-regelingen* (Evaluation of three EZ subsidy tender programmes), 2002, in Dutch

ECN (2000) *Monitoring Energy Efficiency Indicators in The Netherland in 1999. Dutch contribution tot the project 'Cross-country comparison on energy efficiency – Phase 5'*

Fenna, G., *Mure database case study, an analysis of financial measures across the EU*, March Consulting Group, April 1999

Ghalwash, T (2004) *Energy Taxes as a Signaling Device: An Empirical Analysis of Consumer Preferences*, Paper Department of Economics, Umeå University, Sweden

Gibbs, M. and J.C. Townend (2000) *The Role of Rebates in Market Transformation: Friend or Foe?* As published in Proceedings of the ACEEE 2000 Summer Study on Energy Efficiency in Buildings, page 6.121-6.132

Hollomon, B. Ledbetter M. Sandahl,L. and Shoemaker T., *Seven Years Since SERP: Successes and Setbacks in Technology Procurement*, proceedings ACEEE summer study 2002

Ligteringen, J.J. *The feasibility of Dutch environmental policy instruments*, 1999, Twente University, Enschede

Ministry of Spatial Planning, Housing and the Environment (VROM) (2004) *Handreiking voor monitoring en evaluatie van klimaatmaatregelen*

National Renewable Energy Laboratory (1996) *Linking Home Energy Rating Systems with Energy Efficiency Financing: Progress on National and State Programs*, Golden, Colorado

Ostertag, K. *No-regret potentials in Energy conservation*, Physica-Verlag, Technology, innovation and policy serie, number 15, 2003

Rietbergen M, and Blok, K. *The Environmental Performance of Voluntary Agreements on Industrial Energy Efficiency Improvement*. Department of Science, Technology and Society. Utrecht University, Netherlands, 2000.

Taylor, J. and VanDoren, P. (2001) *The Illusion of Energy Efficiency*. www.cato.org/dailys/04-24-01.html

US Department of Energy's Industrial Technologies Platform *Financing Toolbook*, www.eere.energy.gov/industry/financial/fin_toolbook.html

Vermeulen, W.J.V., Das, B.W.J. and Meyer, L.A., *De praktijk van beleidsinstrumenten voor energiebesparing* (The practise of policy instruments for energy savings), TNO, STB/94/006, 1994 (in Dutch)

Violette, D.M. (1995) *Evaluation, Verification, and Performance Measurement of Energy Efficiency Programmes*, USA

Vreuls, H. and van de Laar E, *INDEEP analysis report 2004*, SenterNovem July 2004

Westling, H. *Task X Performance contracting, summary report*, IEA DSM Agreement, May 2003

Westling, H. *Final management report Annex III, Co-operative Procurement of Innovative Technologies for Demand-Side Management*, IEA DSM Agreement, May 2000

World Energy Council (2001) *Energy Efficiency Policies and Indicators. A report by the World Energy Council*, UK

World Energy Council (2004) *Energy Efficiency Policies and Indicators. A report by the World Energy Council*, UK

XENERGY Inc. *Phase 4 market effects study of California residential lighting and appliance program*, April, 2002

Chapter 5: Voluntary Agreements

Cabugueira, M.F.M., *The Voluntary Agreements as an environmental policy instrument in Portugal*, evaluation Proposal, CAVA workingpaper 99/11/13, November 1999

Château, B., *Package of coordinated measures*, Enerdata, presentation ADEME-WEC Workshop on Energy Efficiency Policies,15-16 December 2003,

Chistensen, W. and Espergren, K.A., *Audit II, Topic Report Monitoring and Evaluation*, draft version September 2002

De Clercq M., Seyad A. and, Ameels B. *A comparative study of environmental negotiated agreements*, CEEM, Ghent University, report at the NEAPOL closing conference, December 2001

Eichhammer, W., *Voluntary/Negotiated Agreements for* CO_2 *reduction*, presentation ADEM-WEC workshop on Energy Efficiency Policies, December 2003

Hansen, K. and Larsen, A. *Energy Efficiency in Industry through Voluntary Agreements, the implementation of five voluntary agreement schemes and an Assessment*, AKF, August 1999,

IEA/OECD, *Voluntary Actions for energy-related* CO_2 *abatement*, Policy Analysis Series, 1997

Krarpu, S. and Ramesohl, S., *Voluntary Agreements in Energy Policy, implementation and efficiency*, Final report AKF, January 2000

Minnesota Pollution Control Agency, *Evaluating Voluntary Agreements*, Mercury Appendix D, 2000

Motiva, *Enhanced and Stimulated Evaluation of Energy Efficiency Projects*, 1999, Motiva, *Energy conservation agreements, progress review 2003*, February 2004

Novem, *Handbook energy efficiency monitoring of direct energy consumption in Long-Term Agreements*, February 1999

OECD, *Voluntary Approaches for environmental policies, effectiveness, efficiency and usage in policy mixes*, 2003

Rietbergen, M.G. and Blok K., *Environmental Performance of Voluntary Agreements on Industrial Energy Efficiency Improvement*, Department of Science, Technology and Society
Utrecht University, Utrecht 1999

State Energy Programs (SEP), *SEP Metrics Handbook, Evaluation*, August 2001

Chapter 6: Combinations:

Arentsen, M. *Evaluatie van de MJA-1 en MJA-2 tot en met 2002, Deel A*, CSTM, Universiteit Twente (Evaluation on the LTA-1 and LTA-2 till 2002, Report A) in Dutch, 2004

CLASP website: Collaborative Labelling and Appliance Standards Program = [CLASP S&L Resource Center, General Information on Standards and Labeling – How S&L Work, www.clasponline.org/resource.php3?nnx=5&no=12]

CLASP (Stephen Wiel and Jjames E. MacMahon) 2001, *Energy-Efficiency Labels and Standards: A Guidebook for Appliances, Equipment, and Lighting*, CLASP, Washington DC, USA

DEDE (Department of Alternative Energy Development and Efficiency, Thailand Ministry of Energy) 2002 (Peter du Pont), "*Policy Options for Improving the Energy Efficiency of Electrical Appliances*" - UN ESCAP Ad Hoc Expert Group Meeting on End-Use Energy Efficiency towards Promotion of a Sustainable Energy Future, 18-20 November 2002, Bangkok, Thailand

Dickerson, Ch. A., 2001, *A Framework for Planning and Assessing Publicly Funded Energy Efficiency,* Study ID PG&E-SW040, Pacific Gas and Electric Co. March 2001

Ecofys 2004, *Evaluatie van het klimaaatbeleid in de gebouwde omgeving 1995-2002*, 2004

Egmond, C. The art of changing, behaviour of target groups *(in Dutch:De kunst van het veranderen, gedrag van doelgroepen)* Novem, June 2003

IEA 2000, *Energy labels and standards*

Krarup, S. and Ramesohl, S. *Voluntary Agreements in Energy Policy*, final report from the project VAIE, AKF, 2000

Krarup, S. Wen *Auditing is part of Voluntary Agreements and the risk of Collusion*, AKF, 2001

Neij, L. *Dynamics of Energy Systems, Methods of analysing technology change*, Lund University, 1999

Rogers, E.M. *Diffusion of Innovations*, New York 1995

Ryan, V. *International collaboration on Market Transformation*, final Task VII report IEA DSM Agreement, October 2004

Starzer, O. *Towards Kyoto – Implementation of LTA in industry: which elements make LTA successful and how to integrate them into the policy mix*? Proceedings ECEEE summer study 2001

Starze, O., Betoldi, P. and Sattler, M. *Combining long-term agreements with emissions trading,* Proceedings ECEEE summer study 2003

WEC 2001, EEPI *Energy Efficiency Policies Indicators*

WEC 2004, *Energy Efficiency Policies and Indicators*

Appendix A: Experts participating in the IEA DSM Agreement, Task 1, Subtask 9, Evaluation guidebook

Belgium

Wim de Groote (till January 2004)
Universiteit Gent, Faculteit toegepaste wetenschappen, Vakgroep Mechanica van stroming, warmte en verbranding, Sint-Pietersnieuwstraat 41, 9000 Gent Belgium
Tel + 32 9 264 33 55
Fax + 32 9 264 35 75
Email: Willem.DeGroote@rug.ac.be

Canada

Malika Nanduri, Natural Resources Canada, Office of Energy Efficiency, 580 Boothstreet, Ottawa, Ontaria, K1A EO4, Canada
Tel +1 613 943 2396
Fax +1 613 947 4120
Email mnanduri@nrcan.gc.ca

Denmark

Peter Bach, Danish Energy Authority, Amaliegade 44, 1256 Copenhagen K
Tel +45 33 926700
Fax +45 33 114743
Email pb@ens.dk

Kirsten Dyhr-Mikkelsen, SRC International A/S, Hovedgaden 8, 3460 Birkerød, Denmark
Tel +45 70 20 45 90
Fax +45 70 20 45 91
Email kdm@srci.dk

Richard Schalburg, ELFOR, Rosenoerns Allé 9, 1970 Frederiksberg C, Denmark
Tel +45 35 300 932
Fax +45 35 300 999
Email ris@elfor.dk

France

Didier Bosseboeuf, Ademe, 27, rue Louis Vicat, 75737 Paris Cedex 15, France
Tel +31 1 4765 2355
Fax +31 1 4095 7453
Email didier.bosseboeuf@ademe.fr

Italy

Ornella Celi (till Februari 2004), CESI SFR - Business Unit Usi Finali, Servizi Industria e Rinnovabili, Via Rubattino 54, 20134 Milano, Italy
Tel +39 02 21255299
Fax +39 02 21255626
Email celi@cesi.it

Republic of Korea

Jong-Duck Kim Ph.D., Korea Energy Economics Institute (KEEI)
665-1 Naeson-Dong, Euiwang-Si Kyunggi-Do 437-713, Republic of Korea
Tel +82 31 4202240
Fax +82 31 4202162
Email jdkim@keei.re.kr

The Netherlands (and Operating Agent)

Harry Vreuls, SenterNovem, PO box 17, 6130 AA Sittard, The Netherlands
Tel +31 46 4202258
Fax +31 46 4528260
Email h.vreuls@senternovem.nl

Sweden

Lena Neij, International Institute for Industrial Environmental Economics (IIIEE), Lund University PO box 196 SE-221 00 Lund, Sweden
Tel +46 46 2220268
Fax + 46 46 2220230
Email lena.neij@iiiee.lu.se

Appendix B: Overview of the International Energy Agency (IEA) and the IEA Demand-Side Management Programme

The International Energy Agency

The International Energy Agency (IEA), established in 1974, is an intergovernmental body committed to advancing security of energy supply, economic growth, and environmental sustainability. The policy goals of the IEA include:

- diversity, efficiency, and flexibility within the energy sector,
- the ability to respond promptly and flexibly to energy emergencies,
- environmentally-sustainable provision and use of energy
- development and use of more environmentally-acceptable energy sources,
- improved energy-efficiency,
- research, development and market deployment of new and improved energy technologies, and
- undistorted energy prices
- free and open trade
- cooperation among all energy market participants.

To achieve those goals, the IEA carries out a comprehensive program of energy cooperation and serves as an energy forum for its 26 member counties.

Based in Paris, the IEA is an autonomous entity linked with the Organization for Economic Cooperation and Development (OECD). The main decision-making body is the Governing Board, composed of senior energy officials from each Member Country. A Secretariat, with a staff of energy experts drawn from Member countries and headed by an Executive Director, supports the work of the Governing Board and subordinate bodies.

As part of its program, the IEA provides a framework for more than 40 international collaborative energy research, development and demonstration projects, known as Implementing Agreements, of which the DSM Programme is one. These operate under the IEA's Energy Technolgy Collaboration Programme which is guided by the Committee on Energy Research and Technology (CERT). In addition, five Working Parties (in Energy Efficiency, End Use, Fossil Fuels, Renewable Energy and Fusion Power) monitor the various collaborative energy agreements, identify new areas for cooperation and advise the CERT on policy matters.

IEA Demand Side Management Programme

The Demand-Side Management (DSM) Programme, which was initiated in 1993, deals with a variety of strategies to reduce energy demand. The following 17 member countries and the European Commission have been working to identify and promote opportunities for DSM:

Australia	Italy
Austria	Japan
Belgium	Korea The Netherlands
Canada	Norway
Denmark	Spain
Finland	Sweden
France	United States
Greece	United Kingdom

Programme Vision: In order to create more reliable and more sustainable energy systems and markets, demand side measures should be the first considered and actively incorporated into energy policies and business strategies.

Programme Mission: To deliver to our stakeholders useful information and effective guidance for crafting and implementing DSM policies and measures, as well as technologies and applications that facilitate energy system operations or needed market transformations.

The Programme's work is organised into two clusters:

- The load shape cluster, and
- The load level cluster.

The 'load shape" cluster includes Tasks that seek to impact the shape of the load curve over very short (minutes-hours-day) to longer (days-week-season) time periods. The "load level" cluster includes Tasks that seek to shift the load curve to lower demand levels or shift loads from one energy system to another.

A total of 15 projects or "Tasks" have been initiated since the beginning of the DSM Programme. The overall program is monitored by an Executive Committee consisting of representatives from each contracting party to the Implementing Agreement. The leadership and management of the individual Tasks are the responsibility of Operating Agents. These Tasks and their respective Operating Agents are:

Task 1 International Database on Demand-Side Management &
Evaluation Guidebook on the Impact of DSM and EE for Kyoto's GHG Targets
Harry Vreuls, SenterNovem, the Netherlands

Task 2 Communications Technologies for Demand-Side Management - *Completed*
Richard Formby, EA Technology, United Kingdom

Task 3 Cooperative Procurement of Innovative Technologies for Demand-Side Management – *Completed*
Dr. Hans Westling, Promandat AB, Sweden

Task 4 Development of Improved Methods for Integrating Demand-Side Management into Resource Planning - *Completed*
Grayson Heffner, EPRI, United States

Task 5 Techniques for Implementation of Demand-Side Management Technology in the Marketplace - *Completed*
Juan Comas, FECSA, Spain

Task 6 DSM and Energy Efficiency in Changing Electricity Business Environments – *Completed*
David Crossley, Energy Futures, Australia Pty. Ltd., Australia

Task 7 International Collaboration on Market Transformation- *Completed*
Verney Ryan, BRE, United Kingdom

Task 8 Demand-Side Bidding in a Competitive Electricity Market - *Completed*
Linda Hull, EA Technology Ltd, United Kingdom

Task 9 The Role of Municipalities in a Liberalised System- *Completed*
Martin Cahn, Energie Cites, France

Task 10 Performance Contracting- *Completed*
Dr. Hans Westling, Promandat AB, Sweden

Task 11 Time of Use Pricing and Energy Use for Demand Management Delivery

Richard Formby, EA Technology Ltd, United Kingdom

Task 12 Energy Standards
Frank Pool, New Zealand

Task 13 Demand Response Resources
Ross Malme, Retx, United States

Task 14 White Certificates
Antonio Capozza, CESI, Italy

Task 15 Network Driven DSM
David Crossley, Energy Futures Australia Pty Ltd, Australia

For additional information, see the DSM website: http://dsm.iea.org