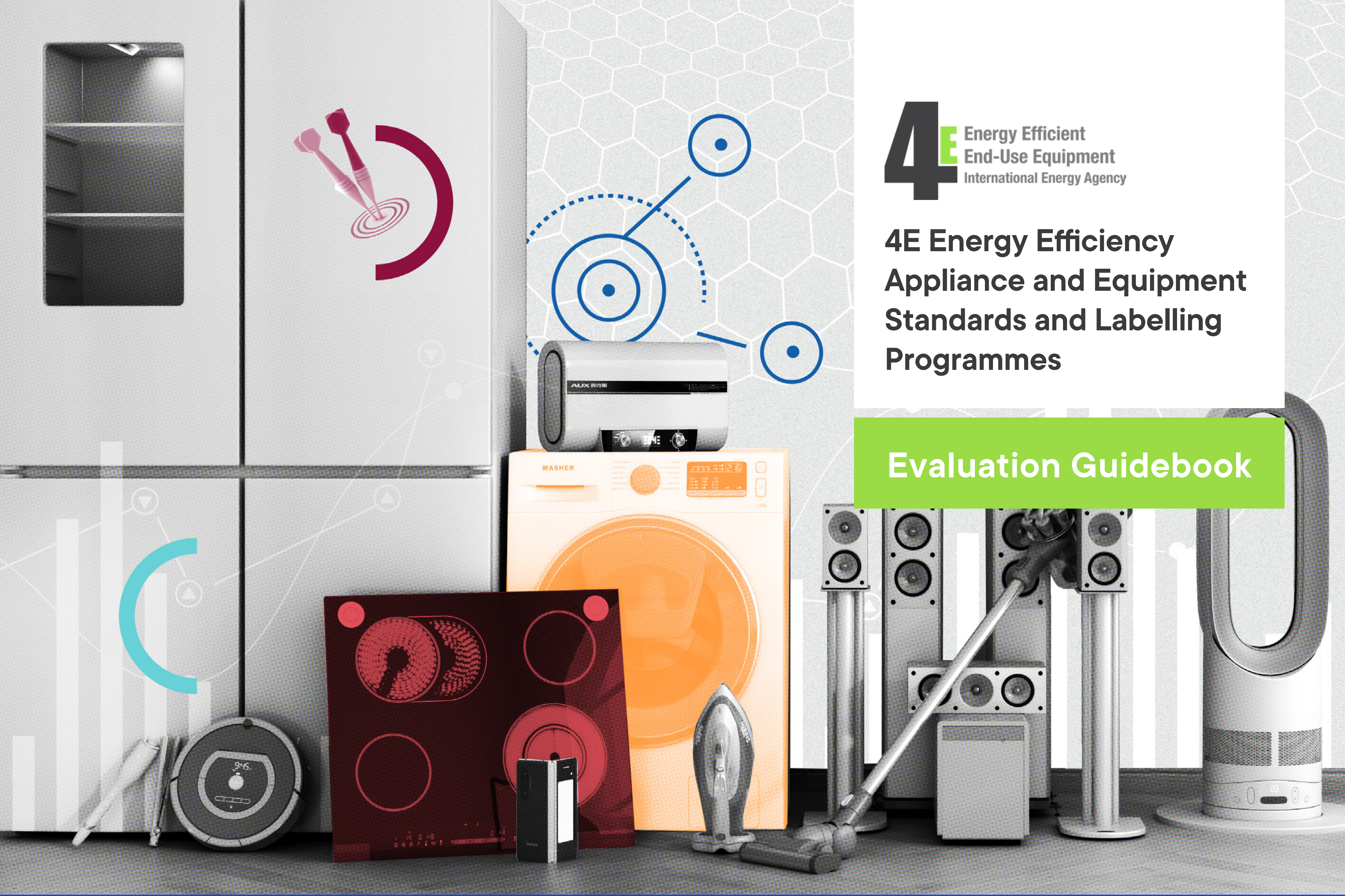


Energy Efficient
End-Use Equipment
International Energy Agency

4E Energy Efficiency Appliance and Equipment Standards and Labelling Programmes

Evaluation Guidebook

Technology Collaboration Programme
by iea

The Technology Collaboration Programme on Energy Efficient End-Use Equipment (4E TCP)

The Technology Collaboration Programme on Energy Efficient End-Use Equipment (4E TCP), has been supporting governments to co-ordinate effective energy efficiency policies since 2008.

Fourteen countries and one region have joined together under the 4E TCP platform to exchange technical and policy information focused on increasing the production and trade in efficient end-use equipment. However the 4E TCP is more than a forum for sharing information: it pools resources and expertise on a wide range of projects designed to meet the policy needs of participating governments. Members of 4E find this an efficient use of scarce funds, which results in outcomes that are far more comprehensive and authoritative than can be achieved by individual jurisdictions.

The 4E TCP is established under the auspices of the International Energy Agency (IEA) as a functionally and legally autonomous body.

Current members of 4E TCP are: Australia, Austria, Canada, China, Denmark, European Commission, France, Japan, Korea, Netherlands, New Zealand, Switzerland, Sweden, UK and USA.

Further information on the 4E TCP is available from: www.iea-4e.org

Acknowledgements

4E TCP wish to thank Dr Lloyd Harrington, Dr Paul Waide and Fiona Brocklehurst for their considerable assistance in preparing the material for this publication.

Contents

Contents

Table of Figures

Glossary & Terminology

The meaning of terms as they are used in this Guidebook are explained here

Additionality	Accounting only for savings, or other changes in attributes, that are in addition to the “without measures” case
APEC	Asia-Pacific Economic Forum
ASEAN	Association of Southeast Asian Nations
Autonomous	Improvement in the performance of equipment energy efficiency not caused by the energy efficiency improvement
Business as Usual (BaU)	Often used to describe the “without measures” or baseline case
Cohort	A group of equipment of the same age
Decomposition	Analysis to determine factors influencing energy use or energy savings
Discount rate	The discount rate is the interest rate used to determine the present value of future cash flows in a discounted cash flow analysis
EES&L	Energy efficiency standards and labelling (programme)
EU	European Union
Ex-ante	Evaluations conducted on programmes before implemented
Ex-post	Evaluations of already implemented programmes
Least lifecycle costs	Assessment of various design options that deliver the lowest lifecycle cost
Lifecycle cost	Assessment of the total cost of a piece of equipment over its lifecycle including initial capital costs, maintenance costs, operating costs and the asset's residual value at the end of its life
Longitudinal trends	Repeated observations of the same variables over a period of time
MEPS	Minimum energy performance standards
METI	Ministry of Economy, Trade and Industry, Japan
Monitoring, Verification and Enforcement (MV&E)	Processes to check for and enforce compliance with obligations under equipment standards and labelling programmes
Non-participants	Individuals or businesses that are not included in a programme
Ownership	The average number of pieces of equipment per household or business. This is effectively the average stock per household/business (usually expressed as a decimal, can be more than 1.0)
Participants	Individuals or businesses that are included, voluntarily or otherwise, in a programme
Penetration	The percentage of households or businesses that have one or more of a particular piece of equipment (maximum 100%)
RFP	Request for Proposal
Rebound	Sometimes called the rebound effect or the take-back effect, it is the reduction in expected gains from new technologies that increase energy efficiency, because of behavioural or other systemic responses. For example, householders heating their homes more after the introduction of more efficient boilers, rather than using less energy
Saturation	The average number of a particular piece of equipment that are installed and in use in homes/business that have the equipment installed (usually decimal, must be ≥ 1.0)
Stock	A total count of a particular piece of equipment that is installed and in use
Tariffs	The charge made for energy supply services
Top Runner	The equipment energy efficiency programme in Japan
UEC	Unit Energy Consumption. Often represents the average annual energy consumption of a product type (or sub-type) for a particular year’s sales under typical or average use. UEC may apply to the average of the sales of new products or the average of the stock of existing products
With measures	The scenario(s) that include a single or multiple policy options
Without measures	The baseline or Business as Usual scenario

Purpose and use of this guidebook

This guide identifies the key steps in the impact evaluation of equipment energy efficiency standards and labelling (EES&L) programmes. These generally include:

- **Minimum Energy Performance Standards (MEPS)**
- **Energy Labelling**

Since evaluations are often undertaken by an independent third party, this guide is designed to provide a clear scope of work or request for proposals (RFP). However, it may also be useful for the general planning of programme evaluations.

Why evaluations are needed

Energy efficiency has been described by the International Energy Agency as the "first fuel" and is the most important strategy to reduce future greenhouse gas emissions (Motherway 2019; International Energy Agency 2021). Energy efficiency has been shown to reduce the overall costs of mitigating carbon emissions while advancing social and economic development, enhancing energy security and quality of life, and creating jobs (Fischer 2021). As the largest potential source of future emissions abatement, it is critical that the impacts of energy efficiency are routinely and accurately quantified.

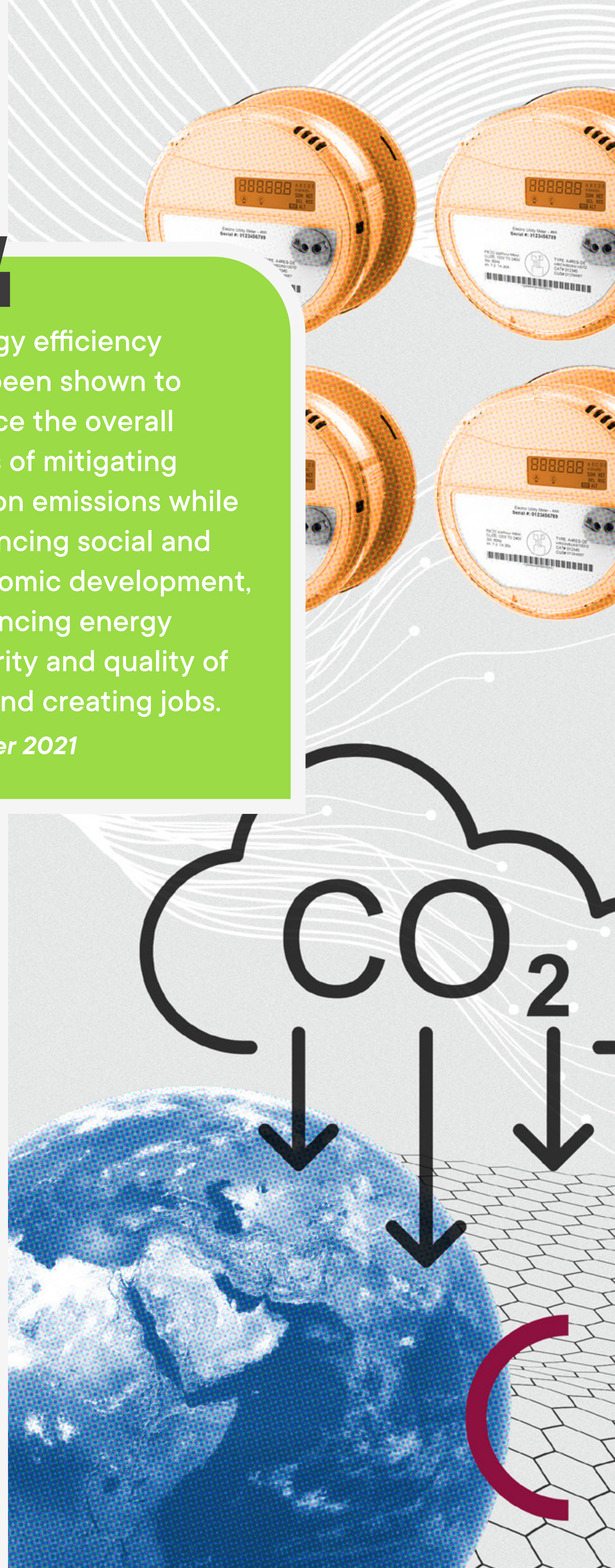
Evaluations of EES&L programmes:

- provide evidence of the impacts, costs and benefits of equipment EES&L policies
- allow the effects of different policies to be compared, both within an economy and internationally[1]
- demonstrate whether policies are working
- identify ways to improve policies
- identify advantages and any disadvantages to particular stakeholders
- allow energy efficiency measures (demand-side) to be compared to energy supply-side options
- enable evidence-based policy choices.

For supply-side energy projects, such as a new power generator, metering the energy output is a relatively simple way of measuring a project's performance. However, this process is not possible for the very large numbers of end-use devices distributed across an economy. Instead, the impacts of energy efficiency programmes are estimated using the techniques highlighted in this Guidebook.

This Guidebook is intended to help generate evaluations that will provide results that are both robust and credible. It draws upon approaches refined over several decades and proven to deliver successful and accurate evaluations of EES&L programmes.

> **Energy efficiency has been shown to reduce the overall costs of mitigating carbon emissions while advancing social and economic development, enhancing energy security and quality of life, and creating jobs.**
>
> ***Fischer 2021***



1 Being able to compare achievements with peer economies is important when determining the success of an EES&L programme and whether it has been sufficiently ambitious.

Content

This Guidebook covers the following five areas:

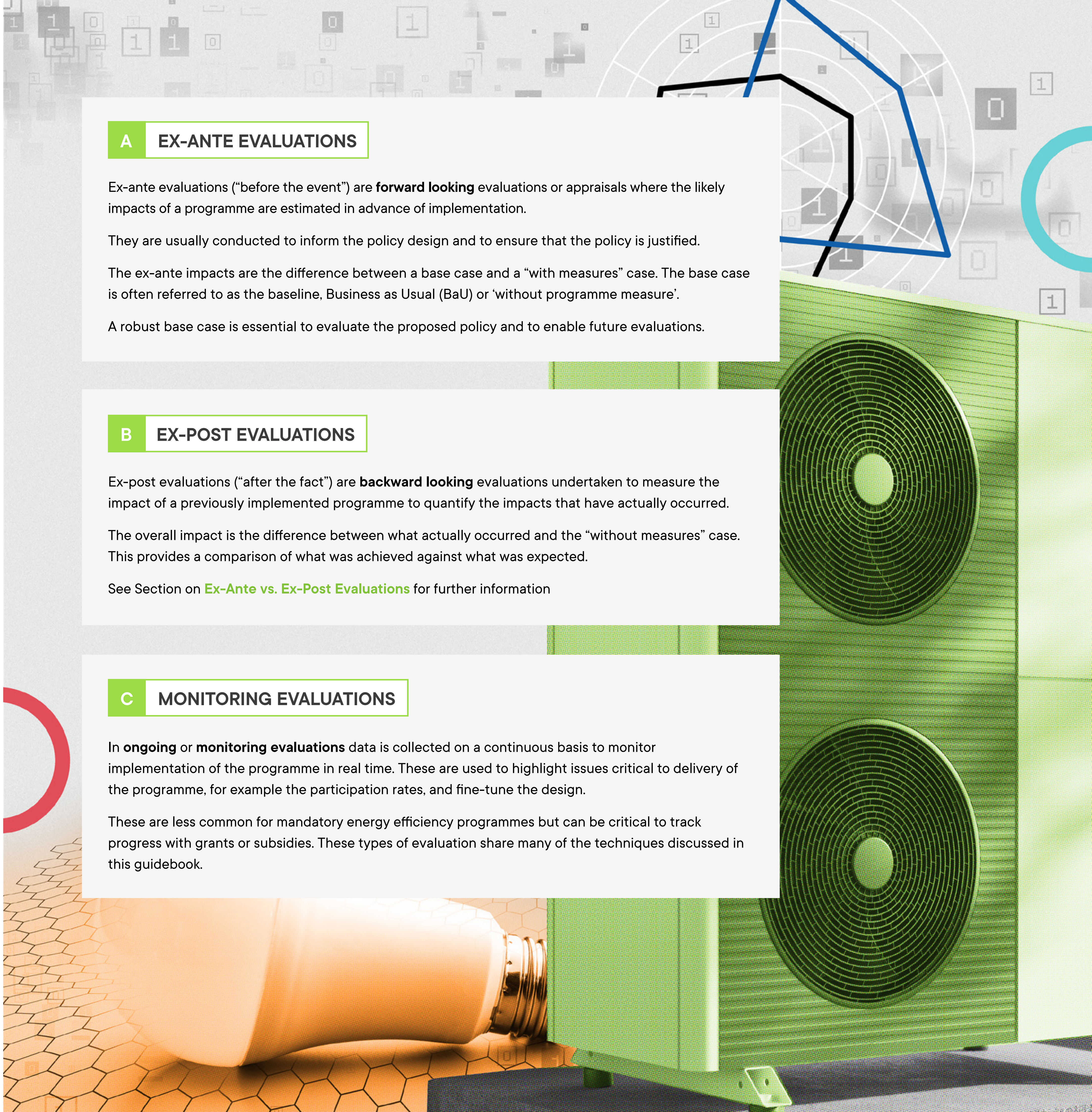
Programme developers can facilitate better evaluations by planning ahead and embedding data collection strategies from the outset.

This Guidebook will help to identify what data will be required and how this can be collected.

Main Types of Evaluation

The most common types of evaluation are described here.

Both ex-ante and ex-post evaluations require comparable data sets and use a similar methodology for analysis, so are closely related.



A EX-ANTE EVALUATIONS

Ex-ante evaluations ("before the event") are **forward looking** evaluations or appraisals where the likely impacts of a programme are estimated in advance of implementation.

They are usually conducted to inform the policy design and to ensure that the policy is justified.

The ex-ante impacts are the difference between a base case and a "with measures" case. The base case is often referred to as the baseline, Business as Usual (BaU) or 'without programme measure'.

A robust base case is essential to evaluate the proposed policy and to enable future evaluations.

B EX-POST EVALUATIONS

Ex-post evaluations ("after the fact") are **backward looking** evaluations undertaken to measure the impact of a previously implemented programme to quantify the impacts that have actually occurred.

The overall impact is the difference between what actually occurred and the "without measures" case. This provides a comparison of what was achieved against what was expected.

See Section on Ex-Ante vs. Ex-Post Evaluations for further information

C MONITORING EVALUATIONS

In **ongoing** or **monitoring evaluations** data is collected on a continuous basis to monitor implementation of the programme in real time. These are used to highlight issues critical to delivery of the programme, for example the participation rates, and fine-tune the design.

These are less common for mandatory energy efficiency programmes but can be critical to track progress with grants or subsidies. These types of evaluation share many of the techniques discussed in this guidebook.

KEY EVALUATION ISSUES

Scope & Objectives

It is good practice for any evaluation to begin by explaining the **policy context and objectives** of the programme under review. The opening section of the RFP will ordinarily:

- set out the background to the policy/programme
- explain its aims and objectives
- summarise its development.

For example, typical programme objectives might include increasing the uptake of energy efficient products in the market to:

- deliver economic and environmental benefits
- reduce health and safety risks and associated societal costs.

The RFP should then set out the **scope of the evaluation** in very broad terms:

- what is its purpose (e.g. to measure the improvements in efficiency)
- what specific products are to be included
- which sectors and market actors are to be considered
- what outcomes are to be assessed and reported (e.g. a set of key performance indicators).

These issues are discussed further in the following sections.

5
CHAPTER

Impact Perspectives

KEY EVALUATION ISSUES
Impact Perspectives

The RFP should specify:

- the actors and perspectives that should be assessed in the evaluation
- the types of impacts to be assessed for each actor

A SPECIFYING THE IMPACT PERSPECTIVES

The perspectives of the following actors are typically considered in EES&L evaluations:

NATIONAL	CONSUMERS	SUPPLIERS	POWER SECTOR	LOCAL INDUSTRY

- the **national**[2] (economy-wide or societal) perspective
- impacts on **consumers** (the buyers and users of the end-use equipment)
- impacts on **suppliers**, potentially including product importers, locally based manufacturers, wholesalers, distributors, physical retail outlets and online
- the potential impact on the **local manufacturing industry**, usually with respect to financial viability and employment impacts
- the effect on the **energy sector** particularly where increasing security of power supply is an important driver for an EES&L policy. If included, the power sector can be split into sub-categories of generation, transmission, distribution and retail.

This Guidebook does not discuss the issue of **demand response capability**, which provides increased flexibility of operation for appliances in response to an excess or shortage of energy on the grid[3].

2 Throughout this guide we refer to the "national" perspective for an evaluation. However, the analysis may be at state/province level, within a utility's service area, or at a multi-country regional level (such as EU, APEC, ASEAN). The term "national" in this guide is intended to mean that the data is aggregated to the geographical level that is required to cover the jurisdiction or customer service area to be addressed by the evaluation.

3 Demand response is primarily a load shifting issue (rather than one of energy efficiency) and these shifts generally either have no effect on total delivered energy consumption, or increase it. But the benefits for energy utilities and consumers are potentially large, so these types of programmes could be evaluated using the principles set out in this guide. This may require some adaptation of the general approaches outlined, with a much stronger focus on energy time of use.

KEY EVALUATION ISSUES

Impact Perspectives

B SPECIFYING THE IMPACT PARAMETERS

The next step is to **specify the impact parameters to be assessed.** Most of these will be quantifiable and will usually cover some or all of the following:

- energy impacts (local and national)
- economic impacts (by each market actor as well as macroeconomic and employment impacts)
- environmental impacts (mostly national)
- social and health impacts (local and national)

The specific impacts to be assessed will depend on which impact perspective is used and type of equipment included. The following sets out the most commonly assessed impacts by market actor.

NATIONAL	CONSUMERS	SUPPLIERS	POWER SECTOR	LOCAL INDUSTRY
> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs	> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs	> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs	> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs	> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs

Different programmes will have different priorities but the elements shown in black text will almost always be assessed.

NATIONAL	CONSUMERS	SUPPLIERS	POWER SECTOR	LOCAL INDUSTRY
> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs	> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs	> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs	> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs	> Energy consumption > Economic e.g. net benefits, cost benefits, balance of payments > Environmental > Health > Employment > Programme costs

If **energy sector impacts** are an important focus, then the impacts on distribution, transmission and generation may be examined. Retail may also be a relevant sub-sector, depending on the market structure.

DISTRIBUTION	TRANSMISSION	GENERATION
> User power demand by hour, day, season, year > Distribution costs by hour, day, season, year > Lost sales revenue per tariff at time of use > User power demand by hour, day, season, year > Avoided marginal investment costs > Security of supply > Employment	> User power demand by hour, day, season, year > Distribution costs by hour, day, season, year > Lost sales revenue per tariff at time of use > User power demand by hour, day, season, year > Avoided marginal investment costs > Security of supply > Employment	> User power demand by hour, day, season, year > Distribution costs by hour, day, season, year > Lost sales revenue per tariff at time of use > User power demand by hour, day, season, year > Avoided marginal investment costs > Security of supply > Employment

If the impact on **domestic manufacturers** is important, then it may be necessary to break this down further.

EXISTING PRODUCTS	NEW PRODUCTS	OVERALL
> Value of lost demand for non-compliant models > Cost of stranded protection assets > Value of demand for compliant models > Revenues > Profitability	> Design and IP costs > Production investment > Material and component costs > Competitiveness > Expected demand > Revenues > Profitability	> Market share/IP/innovation > Revenues > Costs > Profitability

KEY EVALUATION ISSUES

Impact Perspectives

C SPECIFYING FINANCIAL PARAMETERS

The RFP should specify the values used for:

- **discount rate** - a main value and a range to be covered by any sensitivity analysis
- **energy tariffs** – for relevant fuels and sectors, including any real escalation (or de-escalation) factors to account for inflation
- **equipment costs** - over time, including any expected/actual decline (many equipment types show long-term ongoing real reductions in purchase costs over time)
- **environmental costs** - such as real or shadow carbon prices.

In the event that a government agency is commissioning the evaluation, then many of these parameters will already be defined and should be clearly specified.

The RFP may also want to define some **broader socio-economic parameters**, such as job multipliers, or health related parameters if these are in the scope of the evaluation.

D QUANTITATIVE AND QUALITATIVE ASSESSMENT

For each impact to be evaluated, the RFP should state whether the evaluation is to determine **quantified or qualitative impacts**, or both.

Impacts should be quantified wherever feasible.

Qualitative evaluation can be reserved for:

- parameters that are challenging to assess quantitatively such as innovation
- adding explanations to quantitative assessments, such as assessing the role of energy labels in driving the improvement of product performance by industrial actors.

E SPECIFYING THE ANALYSIS TIME HORIZON

The **time horizon,** that is the base year[4] and the end year for the analysis, needs to be stated in the RFP.

The analysis timeframe needs to be such that a reasonable proportion of the stock will have been affected by the programme, taking into account product lifetime and the turnover of the product.

The time horizon may also need to fit with 'standard' timeframes used by government for other greenhouse gas abatement or energy supply measures, to make comparison easy.

4 The base year is the last year where there is zero difference between the "with" and "without" measures cases.

KEY EVALUATION ISSUES

Impact Perspectives

F SPECIFYING THE IMPACT SCENARIOS AND SENSITIVITY CASES FOR EX-ANTE EVALUATIONS

Ex-ante evaluations **require at least one policy scenario** ("with measures") to be considered, but commonly include several variations, such as different levels of ambition and/or timing. These scenarios need to be specified in the RFP.

For example, an RFP may require the impact of different policy options to be compared, and/or different Minimum Energy Performance Standard (MEPS) thresholds to be assessed.

In the case of MEPS, requirements may be set at:

- an energy efficiency level that corresponds to the least lifecycle cost[5] for the consumer
- a level that aligns with existing MEPS in other leading economies
- a level that aligns with MEPS due to come into effect in other economies or corresponding to other programme types.

Where these scenarios rely on parameters that are uncertain at the time of the evaluation and have a significant effect on the impacts (such as energy or carbon prices), then it is good practice to run **sensitivity analyses** for each policy scenario using the mid, upper and lower values.

SUMMARY

Impact Perspectives

- Consider the perspectives and actors to be covered by the evaluation
- For each perspective, define a set of parameters/outcomes to be estimated
- Define the timeframe to be covered
- Identify the policy options to model (ex-ante)

5 The life cycle cost of a product for the consumer is the sum of its purchase price and its discounted operating cost (such as energy purchases) over its lifetime. The least life cycle cost is where this is a minimum.

KEY EVALUATION ISSUES

Methodology & Modelling

A INTRODUCTION

The methodology used in an evaluation has a critical bearing on its accuracy and credibility. However, the most appropriate methodology will vary depending on the number of participants, the nature of the programme and how it is implemented.

While the RFP may provide guidance on preferred approaches and how the analysis could be undertaken, evaluators should be given a degree of freedom to propose their evaluation methodology.

This section provides guidance on the key issues that determine the most suitable methodology and modelling options at a regional or national level, including:

- recommended approaches for **mandatory schemes**
- recommended approaches for **voluntary schemes**
- tracking product **sub-types**
- **adjustments** that may need to be applied, e.g. for weather, changes in use, non-compliance, free-riders, attribution and double counting, rebound effects and actual versus standardised energy consumption
- **impact accounting** via stock modelling to aggregate results
- **financial modelling**.

B CONSIDERATION OF PROGRAMME SCOPE AND TYPE

A key distinction that affects the evaluation methodology is whether the scheme is **voluntary** or **mandatory**.

The major types of programmes and their relevant features include:

- **voluntary schemes** such as incentives, rebates, endorsement labelling and voluntary agreements with supplier groups. These may have a significant number of participants and the effects may vary from weak to strong for participants, but with little to no effect for non-participants.
- **minimum energy performance standards (MEPS)** affect all participants, with moderate to strong impact. This can use engineering or field measurements to estimate the "without measures" and "with measures" cases. Where combined with mandatory labelling, it can be difficult to separately attribute the effect to each programme element.
- **mandatory information schemes**[6] affect all participants usually with weak to moderate impacts when not used with other measures like MEPS. An evaluation needs to employ analytical techniques to estimate the "without measures" and the "with measures" case (especially the underlying autonomous efficiency improvements[7] for the "without measures" case). Because the impact of information is difficult to predict in advance, there is a strong case for ex-post analysis for these types of programmes. This can also help refine future ex-ante estimates.
- **unstructured or general information programmes** may involve a limited or large number of participants (which may be difficult to identify) and/or the overall effect may be generally weak and difficult to quantify, even for participants. These programmes are the most difficult to evaluate accurately.

6 These include mandatory energy labels but also the mandatory provision of product information to consumers.
7 Also called naturally occurring market adoption (NOMAD) in North America.

KEY EVALUATION ISSUES

Methodology & Modelling

C VOLUNTARY SCHEMES

Voluntary schemes include approaches such as incentive schemes, rebates, white certificate schemes and endorsement energy labelling (where only the most efficient products carry a label) e.g. US ENERGY STAR®.

In a voluntary scheme, there will be **participants** and **non-participants**. The methodology selected should aim to establish the effect of participation in the voluntary scheme by comparing both groups.

Direct energy measurement of equipment installed by the participants versus non-participants provides the most robust data. Care needs to be taken in selecting the samples to minimise the influence of factors other than the programme that may influence energy (socio-economic factors, household size, geographic location, etc). Relevant corrections may need to be applied in some cases.

Methods Used to Evaluate Voluntary Programmes

Some of the most common methods used to establish the "without measures" and "with measures" cases are:

- robust sampling[a] of participants and non-participants to establish differences:
  - random allocation/experimental design
  - intervention group ("with measures" case) vs well matched "without measures" case
  - strong difference-in-difference design.
- less rigorous approaches to estimate the impacts for participants and non-participants include:
  - intervention group vs unmatched comparison group
  - predicted vs actual
  - no comparison group.

More detail on these approaches and the associated programme design is contained in supplementary guidance to the UK Government Evaluation Handbook (the Magenta Book) (Campbell & Harper 2012).

a. Robust sampling is where the sample size is sufficiently large for the expected random errors to be much smaller than the expected effect size.

6 These include mandatory energy labels but also the mandatory provision of product information to consumers.
7 Also called naturally occurring market adoption (NOMAD) in North America.

KEY EVALUATION ISSUES

Methodology & Modelling

D MANDATORY SCHEMES

In a mandatory scheme, all end-users are participants, which is why they typically have greater impact than voluntary schemes. It also means that there is no non-participant control group.

Therefore, to estimate the programme impacts, it is necessary to establish longitudinal trends in key parameters, such as energy consumption, product attributes, usage and even purchase behaviour prior to the scheme's implementation.

Changes that differ from the established trends after implementation can then be attributed to the scheme. The better the data on, and understanding of, the market (and trends) prior to implementation[8], the more reliable the calculated impacts will be.

Assessing the impact of MEPS or fleet weighted averages

The lowest possible MEPS impact is assessed when each product type meets the required performance threshold.

However, in reality many suppliers will exceed MEPS, and an evaluation may attempt to calculate what impact this has. The extent that requirements are exceeded depends on how stringent the new requirements are and what cost-effective improvement options are available to manufacturers.

The lowest impact of programmes that are based on fleet weighted averages can similarly be quantified accurately, but again many suppliers may exceed the minimum targets.

This is illustrated in Figure 1 for the Top Runner Programme in Japan.

Mandatory information schemes including labelling

The effect of information, such as mandatory energy labels, is usually more subtle than MEPS and therefore the impacts are often smaller and more difficult to discern.

These programmes aim to help end-users take account of a product's energy efficiency and associated operating costs (amongst other things) when they purchase a new product.

But their effectiveness depends on a range of factors such as the salience[9] of the label itself and the range of reinforcing structures such as online listings, comparison apps and the availability of reliable advice.

The evaluation of information programmes requires a baseline to be established for the trend in standardised energy consumption (as for MEPS). To quantify the impacts of a new programme prior to its implementation requires estimates for the number of participants that will act on the information.

A common challenge for evaluation of information programmes is that a significant minority of end-users will always ignore or be immune to information, no matter how well it is presented.

Periodic consumer surveys, if done consistently, can give insights into purchaser's decision making and highlight the effects of information programmes over time. While these tend to be more qualitative, longitudinal data may be sufficient to establish reasonable quantitative impacts.

Information programmes also aim to increase consumer market pull for more efficient products. As major manufacturers and suppliers are generally highly attuned to consumer demands, they often respond by producing more efficient products where cost effective to do so.

8 Such as trends in share by product sub-type, energy and capacity attributes, price and efficiency.
9 How easy the target audience finds it to understand the information and how relevant they think it is to them.

KEY EVALUATION ISSUES

Methodology & Modelling

An ongoing dialogue with major suppliers can provide some insight into the effectiveness and impacts of information programmes. This qualitative data will help to build up a reasonable basis for a quantitative estimate if regularly collected.

Where both MEPS and information scheme run concurrently, separating the impacts from each other can only be done where there are a range of surveys that provide insight into the contribution of each programme element to the overall savings. This task of allocating impacts to different programmes may be useful to justify or improve some of the programme elements.

Impact Of The Japan's Top Runner Programme For Household Refrigerators

This figure shows the impact of the Top Runner programme in Japan for household refrigerators. Even though suppliers were obliged to meet an index of 100 in the target years, they clearly exceeded these in the target year and continued to improve their products for many years afterwards.

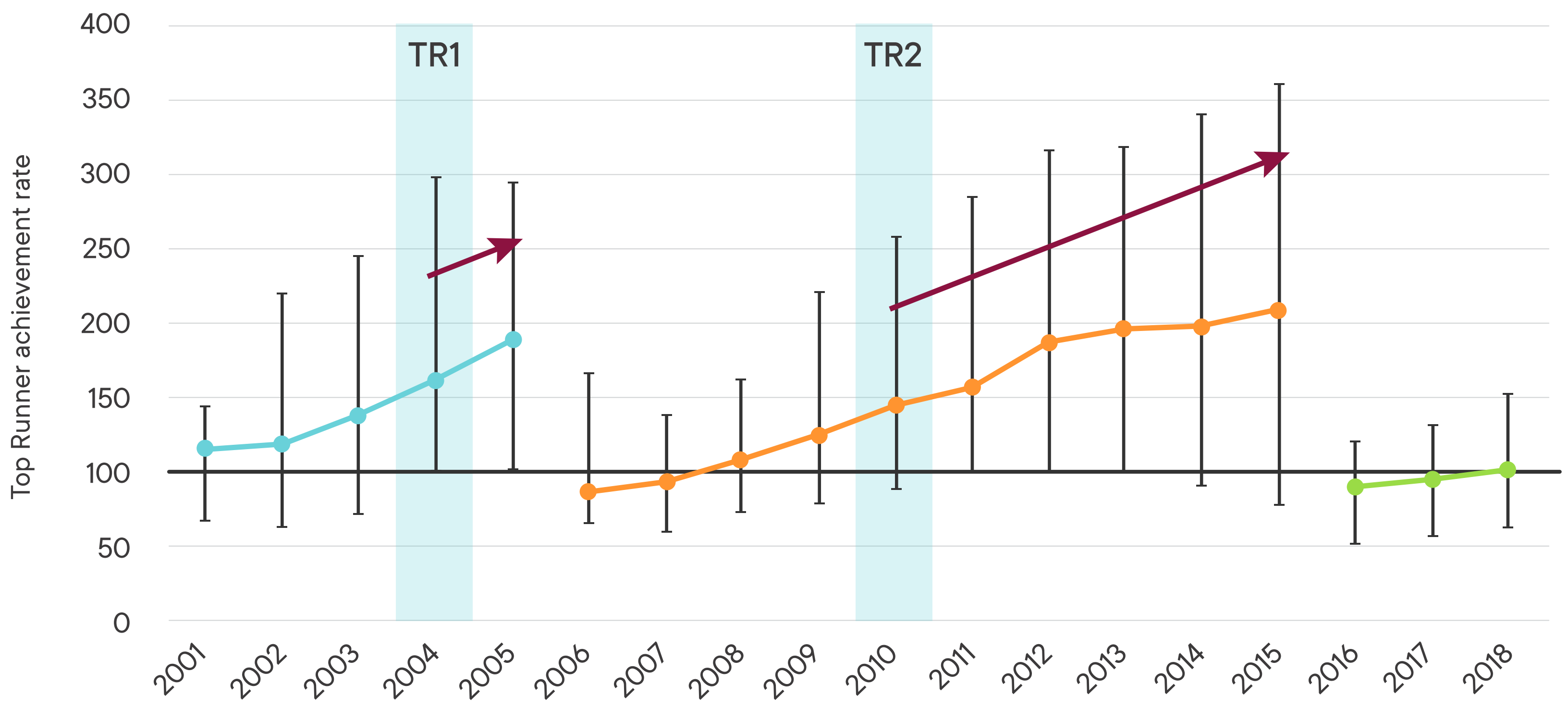


Figure 1: Average standard value achievement rate of new refrigerators each year under Top Runner Programme in Japan

- TR1 = Top Runner target 1 in 2004 , TR2 = Top Runner target 2 in 2010
- Sales weighted average values are shown with error bars indicating highest and lowest performance products each year
- Sales weighted average rate achieved 2001-2005
- Sales weighted average rate achieved 2006-2015
- Market data for the new target set for 2020
- Product efficiency continues to improve markedly even after all suppliers have achieved the Top Runner target
- 100 means that the average product has achieved the specified efficiency target (actual value is reset in each round)

KEY EVALUATION ISSUES

Methodology & Modelling

E EX-ANTE VS. EX-POST EVALUATIONS

Ex-ante evaluations

For forward looking 'ex-ante' evaluations the baseline is established from trends in the available data sets, such as standardised measurements of energy (e.g. shown on an energy label or registered for MEPS).

The "with measures" case then has to be estimated from the available data on the likely impacts at a product level. In the case of MEPS, this can be estimated by examining each product currently available and estimating the energy reduction (if any) that these products would have to achieve in order to comply.

The overall energy reduction (ideally sales weighted) can then be used to establish the "with measures" case. Experience has shown that this is usually a fairly conservative way of estimating energy impacts, but it is robust and defendable.

Ex-post evaluations

'Ex-post' evaluations aim to quantify the impacts of a programme that is already implemented.

In this case, the actual market is the "with measures" case, while the original 'ex-ante' evaluation can be used, with adjustments, as the "without measures" baseline.

To get an accurate estimate of the 'ex-post' impact of the programme, the original "without measures" case needs to be adjusted to take into account differences between the assumptions used in the original ex-ante projection and what actually transpired. These adjustments may
include, for example, the number of households; the ownership rate; the share of each product sub-group and the likely energy characteristics of each sub-group.

In reporting the results, analysis can be used to better understand what impact the changes in each original assumption had on the overall savings estimate (see Figure 2).

Example of Ex-Post Analysis

This figure shows the difference between the forecasted savings from the original ex-ante evaluation (on the left) and the actual savings ex-post shown (on the right). In this case, the difference between the predicted and actual savings were caused by underestimates in the number of households, ownership and product size. However, the original evaluation slightly overestimated the size of freezers. The other changes shown were mostly associated with larger than predicted reductions in energy consumption as a result of the programme.

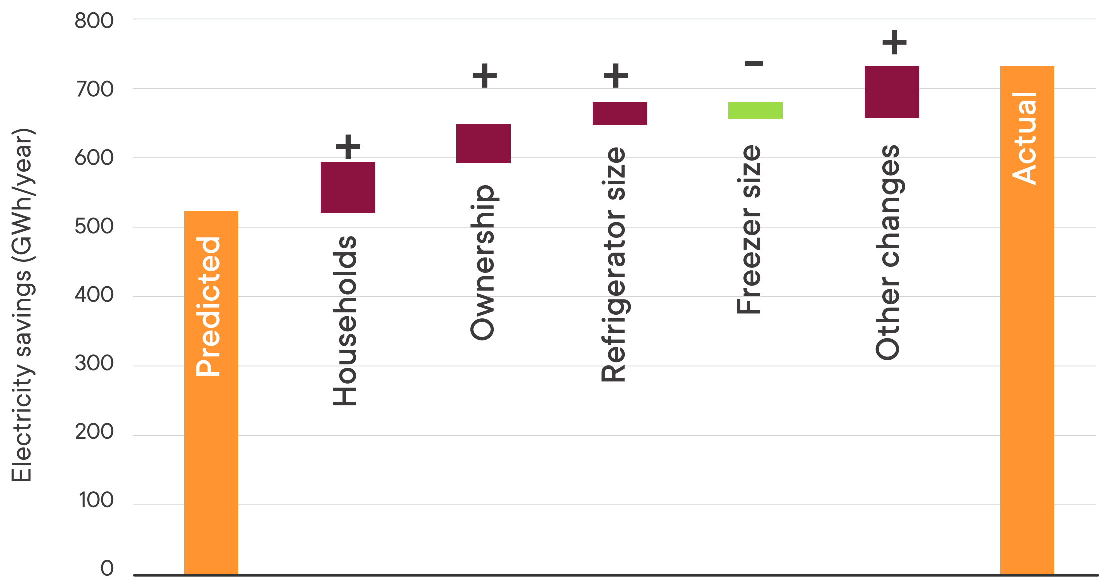


Figure 2: Changes in 2009 energy savings from MEPS-2005 for refrigeration appliances - Australia

Source: Figure 16 from Harrington and Lane (2010)

KEY EVALUATION ISSUES

Methodology & Modelling

F TRACKING PRODUCT SUB-TYPES

For major product categories, like household refrigerators or water heaters, there may be up to 10 or more sub-types that have different characteristics and therefore need to be monitored and modelled separately. If the share of larger, or more energy intensive sub-types increases over time, this means that the overall energy consumption of the product category will also increase, even when there is no change to the energy efficiency within each sub-type. In effect, this means that a single product type may require multiple sub-models in order to get the most accurate picture of baseline trends and potential programme impacts.

G OTHER FACTORS TO CONSIDER

In order to develop robust estimates of energy impacts a number of corrections may need to be applied to adjust for external influences, sampling or participation. Some of these are set out below. The RFP will need to specify which, if any, of these corrections need to be included in the evaluation.

Issues that need to be considered in the "without measures" and "with measures" cases can include:

- **normalising for the impact of weather** - this mainly applies to heating and cooling appliances in an ex-post evaluation when interpreting end use measurement data
- **normalising for the impacts of other changes in usage** – i.e. taking account of any programme induced changes in user interaction or behaviour. This may require examining the underlying trends in usage prior to the programme implementation e.g. hours watching a TV, hours of light use, number of cycles per week for clothes washers, dishwashers, dryers, etc. There may also be secondary usage impacts from changes in appliance capacity: e.g. if new clothes washers have a larger capacity, users may do fewer loads per week or the same number of loads at smaller percentage of rated capacity. This may significantly impact energy use.
- **adjustments for participation** - including free riders, i.e. people that benefit from the programme who would have selected the more efficient product in any case; and spill over effects i.e. where non-participants adopt the measure due to programme influence.
- **additionality** - only counting savings that are in addition to the "without measures" case, after taking account of autonomous efficiency improvement.
- **double counting** – ensuring that savings from complementary measures are not inadvertently counted twice.
- **non-compliance** - the 'with measures' case usually assumed 100% compliance, however where the Monitoring, Verification and Enforcement (MV&E) regime is insufficient, savings may be overestimated.
- **direct rebound** - the situation where demand for an energy service increases if the efficiency, and therefore cost, of that energy service becomes cheaper. This is most likely to occur in cases where energy services are constrained due to their cost, affecting their affordability, such as heating in low-income households[10].
- **indirect rebound** - where the demand for discretionary services (such as travel or entertainment) increases if the efficiency, and therefore cost, of energy services becomes cheaper. This tends to be a macroeconomic issue.

Another important consideration is the determination of **actual energy usage**, as opposed to energy use under standard test conditions. If standardised energy consumption data is available e.g. estimates of energy consumption on an energy label or in an energy database derived from measurement using a standard test procedure, it may be necessary to convert this to the corresponding value of typical energy consumption during normal use.

This may mean that the energy characteristics and/or usage need to be changed. There will always be a spread of usage, so having some understanding of the distribution of normal usage is helpful when attempting to identify groups of winners and losers from a programme.

10 Note that if a direct-rebound effect occurs then the level of service has increased, even if net energy savings are lower than expected.

KEY EVALUATION ISSUES

Methodology & Modelling

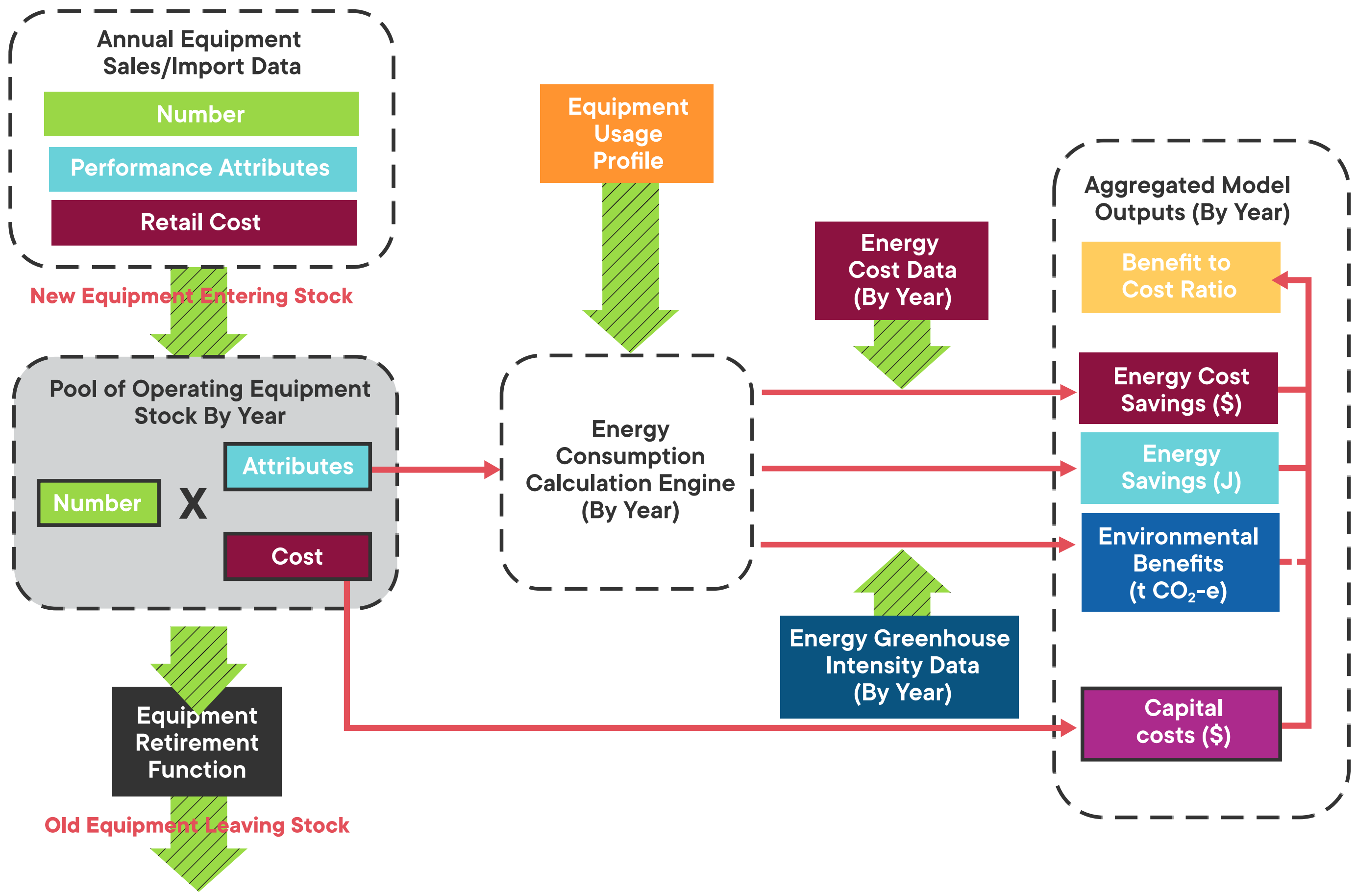
H STOCK MODELS TO AGGREGATE IMPACTS

For appliances and equipment, the number of new products sold and entering into use is often thousands or even millions per year, so a model will be needed to quantify and aggregate a programme's impact.

Stock models provide a convenient way of calculating equipment energy use, product costs, energy bills, emissions and other key parameters for the whole stock of installed products.

Stock models also provide a number of other functions, such as allowing adjustments to changes in ownership, the total number of households or businesses over time, and keeping track of currently in-use products and retirement of older products from the stock at the end of their lifetime. Figure 3 gives a schematic description of a comprehensive stock model.

Figure 3: Schematic of a typical comprehensive stock model



There are a wide range of stock models available, both proprietary and open source, varying in their functionality. However, stock models can also be created in relatively simple spreadsheets.

The RFP should specify the requirements of a stock model and should invite the proposers to describe their proposed approach to stock modelling, including its ability to track the key attributes.

Requirements may also specify the ability to scrutinise the data inputs, underlying calculations and data outputs.

Suggested requirements for stock models used to conduct a quantified impact assessment for EES&L programmes are set out in Annex A.

KEY EVALUATION ISSUES

Methodology & Modelling

I FINANCIAL MODELLING

Aggregated outputs of energy consumption, emissions, product costs and energy costs can be used to perform a financial analysis for each case and compare the overall results. In order to compare financial streams fairly, it is usual to calculate the Net Present Value (NPV) of each case.

NPV is usually present in spreadsheets and other software and is represented by the following formula:

$$NPV=\sum_{t=0}^{n}\frac{R_t}{(1+i)^t}$$

WHERE:

NPV is the Net Present Value of the cash stream being analysed

Rt is the cash inflow and/or outflow that occurs in year *t*

i is the discount rate

t is the time period (year) of the cash flow

n is the total number of periods (years) in the series.

The discount rate (and sensitivity range) is usually specified in the RFP and is often dictated by government Treasury Departments. The discount rate is very important in financial evaluations; a higher discount rate gives higher value to up-front costs and lower value to future operating costs. For long term public goods such as mandatory EES&L programmes, there is a strong argument to use low discount rates (<5%).

In most cases, the stock model can be used for financial impact analysis, but in some cases more specialist financial models may be required (e.g. macroeconomic impacts, employment, health).

SUMMARY

Methodology & Modelling

- Indicate the preferred evaluation methodology and invite detailed proposals
- Specify the required capabilities of a model
- Identify which parameters are to be covered
- Identify any sensitivities to be analysed and specify the range
- Specify which product sub-types need to be treated distinctly
- Specify the type of financial modelling to be used

KEY EVALUATION ISSUES:
Data Input

A INTRODUCTION

The RFP should specify the **minimum data inputs** that will be needed to conduct the evaluation. It should indicate:

- any known datasets or sources that can be used as inputs
- where there are no readily available sources, bidders should be invited to propose their approach to gathering the required data (and include the cost).

The key information that needs to be collated as a time series for the “with measures” and “without measures" cases in the evaluation includes:

- product energy attributes - the energy consumption and efficiency characteristics of new products
- product price attributes - the purchase price and the relationship between price, energy efficiency and energy consumption
- stock flows and turnover - annual sales, effective product lifetime[11], stock in any particular year, along with stock characteristics (i.e. split by size and type)
- changes in product use and capacity over time.

Consideration should be given to obtaining critical input data before commissioning the evaluation. Alternatively, governments may be able to help evaluators access data by making the request directly or providing evaluators with letters of support.

Private sector data is often confidential and of commercial value, so mechanisms may be needed to use the data in a way that respects that confidentiality. This could be via formal or informal confidentiality agreements and commitments to only publish aggregated data[12].

B PRODUCT ENERGY ATTRIBUTES

Data needs

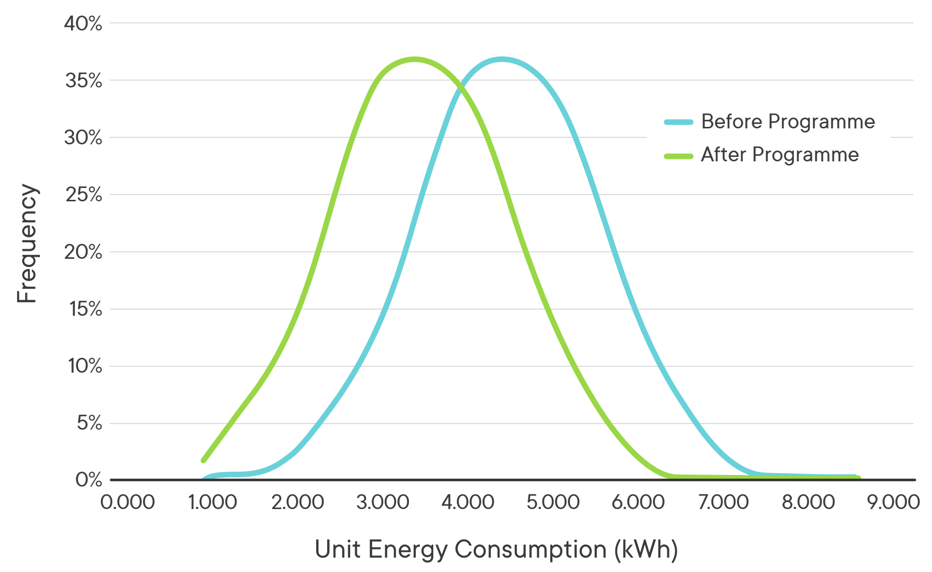
EES&L programmes typically focus on the energy characteristics of **new products.**

An evaluation needs to establish the baseline situation (before the programme) **and** the effect of the policy option.

The core parameter that lies at the heart of any EES&L evaluation is the **unit energy consumption (UEC[13])** of each new product.

The UEC is the average annual energy consumption of the product type (or sub-type) for a particular year's sales under conditions of normal use[13]. Ideally, the UECs will be a series of distributions. A typical example of distribution of product energy efficiency is shown in Figure 4, in this case for both before and after the programme.

Figure 4: An example of product energy efficiency before and after the introduction of energy labelling



11 This may be the same as the technical product lifetime, but for some products this is shorter if products are replaced before they fail.
12 Most commercial organisations will require confidentiality agreements as a condition of purchase of such valuable data, so it is important to clarify any intended use beforehand.
13 The UEC of existing products in the stock may also be useful. Other average product characteristics, such as size or capacity, may be tracked in a similar manner to the UEC.

KEY EVALUATION ISSUES:
Data Input

A schematic on how UEC is calculated is illustrated in Figure 5.

The UEC multiplied by the stock gives the overall energy consumption for the cases with and without the policy measure. From this, other data such as emissions can be calculated. The costs for equipment users can be calculated from the expected/actual equipment purchase costs, the UEC, data on energy tariffs and other relevant consumables (such as water). Further information about the calculations undertaken by a stock model are provided in Annex A. Other non-energy indicators flow from these aggregated parameters.

Thus, the core of the evaluation is the change in the new product UEC resulting from the policy and how this affects the energy consumption of the stock.

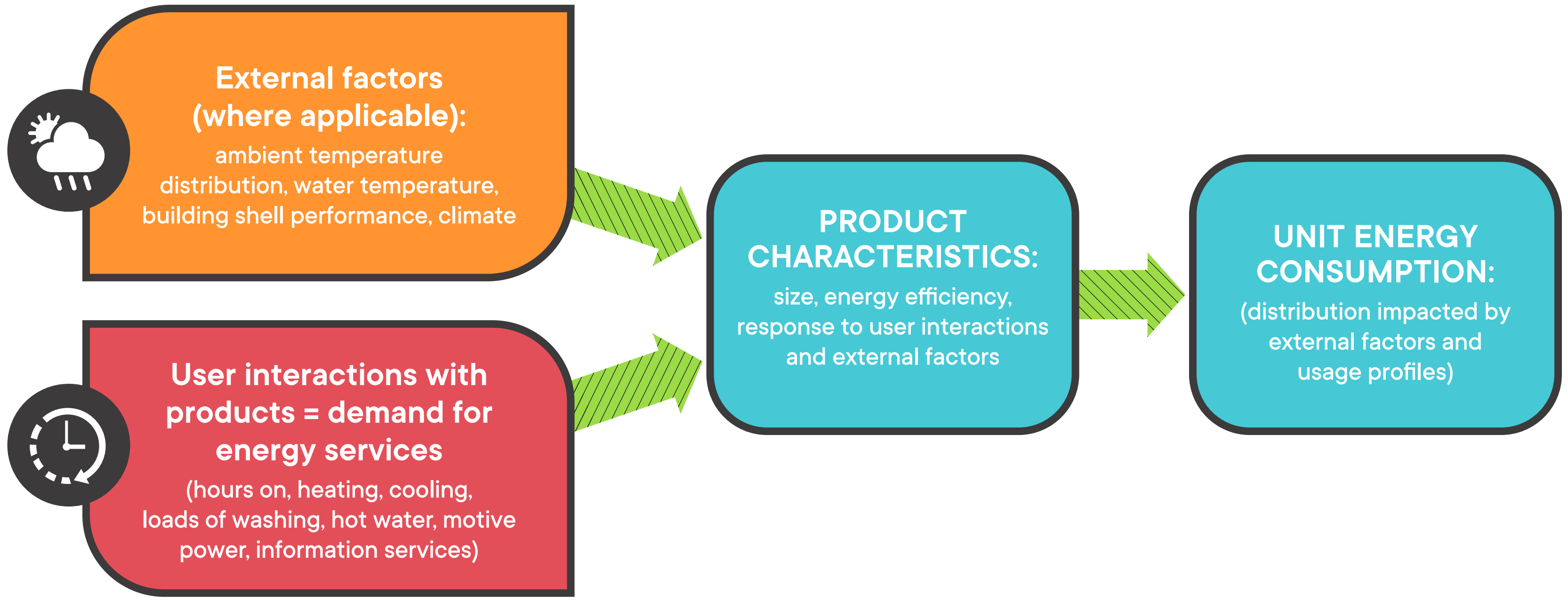
Data sources

If an EES&L programme has a **product registration database**, this will contain data for each of the key product technical and energy performance characteristics.

Other alternatives sources include:

- voluntary product certification schemes often include public databases covering a significant proportion of the products sold in a market
- industry or manufacturer association databases of members' products
- market surveys, including enquiries to each supplier about their product range
- web crawling to gather details from online sources (Mogensen et al. 2019)
- consumer association databases
- commercial market research agency databases
- crowdsourcing, including individuals recording key information on products from retail outlets.

Figure 5: Schematic of the calculation of Unit Energy Consumption (UEC)



KEY EVALUATION ISSUES:
Data Input

C STOCK, SALES AND OWNERSHIP

Data needs

Usually sales, stock and ownership are the most important parameters to track as they are a count of the number of units that are installed and consuming energy. Stock may also be derived from historic sales data.

Where information on the sales volume of specific models is not available, it is common practice to assume that each model known to be on the market sells equally[14] or to weight model sales by the market share of that supplier.

For markets dominated by imports, assessing the quantity and value of imports by country of origin can help to calibrate the estimated share of brands originating from those countries.

Where registration databases are combined with market data, such as brand share or model sales, they can give an accurate weighted average result.

See Annex A for further details on Stock Modelling

Data sources

The number of historical and projected households or businesses are a good basis for making future stock estimates. These are useful in quantifying the total market size and broad breakdown.

Data on market volumes, values and shares can be acquired from:

- **commercial market data** and reports
- **surveys of suppliers** (manufacturers, importers, but also distributors and retailers
- **processing of customs data** (value, volume, weight) on imports and exports
- Some **statistical agencies** also collect data on locally manufactured products

Information of product ownership can come from:

- government surveys and statistics
- utility surveys and statistics
- consumer surveys and interviews
- end use monitoring data (where individual products are metered in situ)
- commercial market research organisations.

Where sources only cover some historical years, data may need to be interpolated. Often it is easier to project trends based on ownership levels than on stock numbers or inferring from sales, since the change in ownership is usually slow. Depending on whether the product is essential, desirable or obsolete, ownership may increase, remain stable or decrease over time.

14 Model weighted average market data are useful where there are a large number of models (>100) and where the market share of any particular model is not large. This approach is less accurate where there are a small number of models or where a few models dominate the market.

KEY EVALUATION ISSUES:
Data Input

D PRODUCT PRICE ATTRIBUTES

Data needs

Product price data is important to quantify the **cost of providing the product**, or its energy service, over the programme timeframe.

To determine the programme impact on product prices, evaluations also have to examine the **relationship between product energy consumption and product price**. This relationship is central to estimating the cost-benefit impacts of an energy policy and requires sales and price data for most of the market in a specific year (ideally the price actually paid rather than the advertised price).

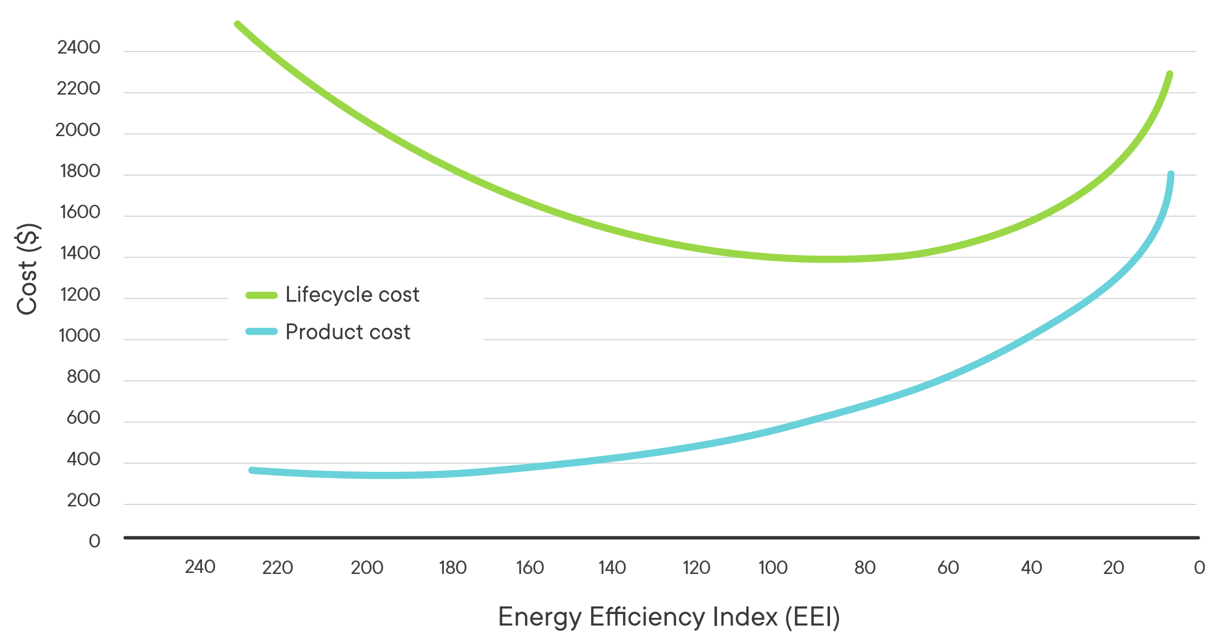
This techno-economic energy engineering analysis considers how successively higher efficiency design options will increase product costs (based on the cost of components and manufacturing) and how the resulting increased efficiency will reduce running costs. Figure 6 shows an example of a resulting curve of product price and life cycle cost as a function of energy efficiency.

In some instances, it may be appropriate to adapt analyses produced in another economy to reflect the local market conditions, rather than attempting to derive all the values from first principles (which tends to require substantial amounts of confidential and/or proprietary data as well as significant effort).

While product costs as a function of efficiency are often relatively stable across jurisdictions, their usage, energy costs, duties & taxes, and supply chain mark-ups may vary considerably.

It is important to note that the product price paid by the final purchaser includes the cost of manufacture, transportation, and mark-ups for the manufacturer, importer/wholesaler and retailer, as well as any customs duties and sales taxes. Understanding the breakdown of these costs can be used to project the expected change at each point in the supply chain and hence how each market actor may be affected.

Figure 6: Relationship between product price, lifecycle cost and energy efficiency



Note: For the efficiency index, a lower value means a lower energy consumption (higher efficiency) relative to a reference product.

Data sources

Product price data can be acquired from the same sources as described in the previous section.

Detailed energy attributes matched with the actual retail price paid will yield the most accurate data.

KEY EVALUATION ISSUES:

Data Input

The Relationship between Product Price and Energy

Traditional economic theory suggests that more efficient products should cost more due to the inclusion of higher quality components and materials. However, many detailed studies using comprehensive data sets have only found a weak correlation between purchase price and energy performance (Energy Efficient Strategies 2016; Energy Efficient Strategies et al. 2021; ENERVEE 2014; EA/4E TCP 2021).

In addition, many research papers show that real (inflation adjusted) prices of many products subjected to energy efficiency standards and labelling are falling in real terms over time, even when their energy consumption is also falling (Ellis et al. 2007; Energy Efficient Strategies 2016; Harrington 2017; Energy Efficient Strategies et al. 2021; Weiss et al. 2010; Desroches 2013; US Department of Energy 2011).

It appears that there are a number of factors that can affect product purchase price, with energy performance just one of many factors, which sometimes seems to have a small impact.

E USAGE PARAMETERS

Data needs

Knowing how and when products are used is critical. For example, the wash temperature, the number of loads per week and the capacity utilisation of a clothes washer are needed to accurately calculate its energy and water consumption.

Similarly, the hours per week that a television is on will be one of the main drivers of its energy consumption.

It is important to understand that changes in usage will not affect just new products in a particular year but usually all of the stock equally.

Data sources

Data on usage can be obtained from:

- government surveys and statistics
- private data collection and statistics (e.g. organisations that collect data on television watching habits for commercial rating statistics)
- utility surveys and statistics
- consumer surveys and interviews
- end use monitoring data (where individual products are metered in situ).

User operation of space conditioning appliances (heating and cooling) will be driven by climate, but also influenced by building shell performance, occupancy and zoning. Lighting use may also be strongly seasonal, depending on latitude.

When compiling data on different types of end-use equipment, it is useful to group products by those whose energy consumption is:

- **strongly driven by user behaviour:** televisions, computers, clothes washers, dishwashers, clothes dryers, hot water demand, lighting, motors
- **moderately affected by user behaviour and climate:** refrigeration systems, which can be impacted by season and climate and modestly by direct user interactions
- **relatively insensitive to user behaviour:** broadband and network equipment, heat losses for hot water storage, distribution transformers
- **complex products:** space conditioning and solar water heating, which depend on climate/weather, equipment capability, equipment efficiency and user interactions.

KEY EVALUATION ISSUES:
Data Input

F
AREAS OF SPECIAL CONSIDERATION REGARDING DATA INPUTS AND ASSUMPTIONS

There are some areas where particular care is required to ensure that the evaluator deals with data and outputs in a consistent and appropriate manner. This may be addressed by the inclusion of relevant components in the specification, as applicable.

Some examples where the RFP may need to specify particular approaches may include:

- assessing the impact of the programme on equipment costs
- where usage is changing over time, a scaling factor may need to be added
- converting energy consumption measured under standard test conditions to energy consumption or performance during normal use (via the application of climatic and usage adjustment factors where applicable)[15]
- estimating equipment lifetime (and its distribution)
- estimating sales and stock (see Section on **Stock, Sales and Ownership**)
- dealing with equipment that is incorporated into other products (such as motors in washing machines)
- accounting for variations across different sectors, for example in usage patterns[16]
- assessing the programme impact on equipment size/capacity, usage or other performance aspects[17]
- estimating programme implementation and operation costs.

SUMMARY

Data Input

- The quality of data used in the evaluation will critically influence the accuracy of the evaluation
- Specify minimum data inputs and potential sources
- Consider purchasing relevant data before the evaluation
- Specify the policy options and sensitivities to be evaluated

15 MEPS and labels are based on energy consumption measured under standard conditions, whereas the energy savings materialise under real life conditions. Where a measurement standard does not reflect average usage conditions, this should be taken into account in the evaluation of actual energy savings.
16 Product usage profiles will often vary in a systematic way depending on the sector in which they are used. For example, split room air conditioners can be used in both the residential and commercial sectors, but the number of hours they are used, the time of day they are used and the loading can vary substantially between these sectors.
17 Most programmes assume capacity and usage are not directly affected by changes in energy or efficiency, unless there is a good reason to think otherwise, but some programmes aim to influence user behaviour.

KEY EVALUATION ISSUES:
Reporting

A OVERVIEW

This section concerns both the provision of data to the client and the public reporting of results. It suggests some standardised best practice approaches and a range of issues to consider when specifying reporting in an RFP.

Templates are useful to ensure that the programme manager obtains the essential information in a format that is useful. This applies to input data, including any assumptions, as well as the results (output data). Examples of templates for input and output data are included in Annex B of this guide.

For complex evaluations, the RFP may specify that the output is provided as an annotated spreadsheet. This makes it easier to present the data in different ways and is helpful in considering the sensitivity of results to changes in key factors.

All data should be presented in consistent and internationally recognised units, ideally SI units[18]. Financial information should normally be presented in the local currency.

The results should clearly state if primary or delivered energy values are used. Where values are converted from primary to delivered energy the conversion factor should be stated, noting that this may vary over the evaluation timeframe.

In **published reports**, evaluation results need to be clearly presented to facilitate their interpretation and maximise their credibility. Best practice is to publish as much of the underlying data and assumptions as possible, however, the amount of data presented will depend on the audience. It may be appropriate to aggregate data for similar or related products together.

It is important for the transparency of the evaluation that the input data are clearly described and referenced and assumptions are stated and justified. This should include documenting any feedback from stakeholders that was recorded during stakeholder consultation.

Particular issues and areas of uncertainty with parameters should be identified, for example if a data source is somewhat dated, but no more recent and robust source could be identified. Ideally a confidence level should be assigned to each data set, based on data quantity and source integrity/robustness of data.

B PROVIDING DATA INPUTS

The key inputs include any historic data used as well as externally defined variables and data that are specified over the time frame of the evaluation.

Input data should be provided with references where appropriate and/or the rationale underpinning any assumptions. The range of data provided may include:

- electricity and fuel prices
- **marginal CO_2-e emission intensity factor** for electricity or other fuels (noting that these may change over time with decarbonisation)
- carbon or shadow **carbon prices**
- the discount rates and time horizon to be used
- emissions that affect **air quality** (SOx, NOx etc.) for each fuel and if their impact is local, e.g. from a biomass boiler, or national, e.g. electricity generation
- activity drivers such as **number of households**, or **commercial floor area**
- administrative costs to government

18 There may be exceptions to this - some jurisdictions may use non SI units locally for the specific technology e.g. Coefficient of Performance (COP) for the energy efficiency of air-conditioners. Where it is more appropriate for non SI units to be used, for example to reflect local usage norms, conversion factors should be provided or alternative tables with SI units should also be provided to facilitate international comparisons.

KEY EVALUATION ISSUES:
Reporting

- administrative costs to manufacturers
- conversion factor from primary to delivered energy (if used)
- rate of inflation (if used)
- assumption on compliance rate (if used).

Common input data that will be used in most analyses and should be provided for each product are:

- stock numbers (directly or from % ownership by household or for businesses)
- data on sales by sector and sub-type where applicable
- key usage parameters such as hours of use or type of usage
- lifespan expressed in hours used or elapsed time in years
- product capacity, size or volume
- product costs for consumers in the "without measures" and "with measures" cases
- energy consumption (by fuel type) per product in the "without measures" and "with measures" cases
- additional compliance costs to manufacturers
- the start date and performance requirements of the policy.
- key usage parameters such as hours of use or type of usage
- lifespan expressed in hours used or elapsed time in years
- product capacity, size or volume
- product costs for consumers in the "without measures" and "with measures" cases
- energy consumption (by fuel type) per product in the "without measures" and "with measures" cases
- additional compliance costs to manufacturers
- the start date and performance requirements of the policy.

C REPORTING OUTPUTS

The RFP should clearly state which outputs are required and how this should be presented. These should be related to the aims of the evaluation, as described in Scope, aim and objectives of the evaluation.

Energy-Related Outputs

The main reported energy-related outputs, provided in tables and graphically, normally include:

- national: energy consumption and savings by fuel type and in total by year for each sub-model covered
- national: environmental: CO_2-e emissions, other emissions by year
- national: changes in product energy consumption or efficiency over time
- consumers: changes in energy and other consumables bills by year over the evaluation timeframe
- consumers: changes in equipment capital costs by year over the evaluation timeframe.

KEY EVALUATION ISSUES:

Reporting

Additional variations can include:

- impacts from the perspective of different market actors (described in the section on **Impact perspectives & parameters**)
- sensitivity of impacts under different input parameters (e.g. growth rates, discount rates). When included, the justification for selecting these cases should be explained.

To give the results context the total values of parameters for the "without measures" and "with measures" scenarios as well as the savings should be reported. For example, an energy saving of 2TWh/year could be substantial or trivial, depending on the "without measures" energy use. This is particularly important when comparing relative impacts and effectiveness between different countries and regions.

Cumulative data, such as total energy savings over the evaluation timeframe period, are not a substitute for reporting year on year values. Cumulative totals should always specify the relevant year ranges covered.

Where the evaluation is to meet a statutory requirement, e.g. a Regulatory Impact Assessment, the presentation of the results may be tightly prescribed.

It may be beneficial to request additional information to ensure transparency and to cover future eventualities, including future ex-post evaluation and international benchmarking.

If the evaluation includes stakeholder consultation, then the process of how this was undertaken, the results and how stakeholders' views were taken into account in the evaluation should be documented.

Financial Outputs

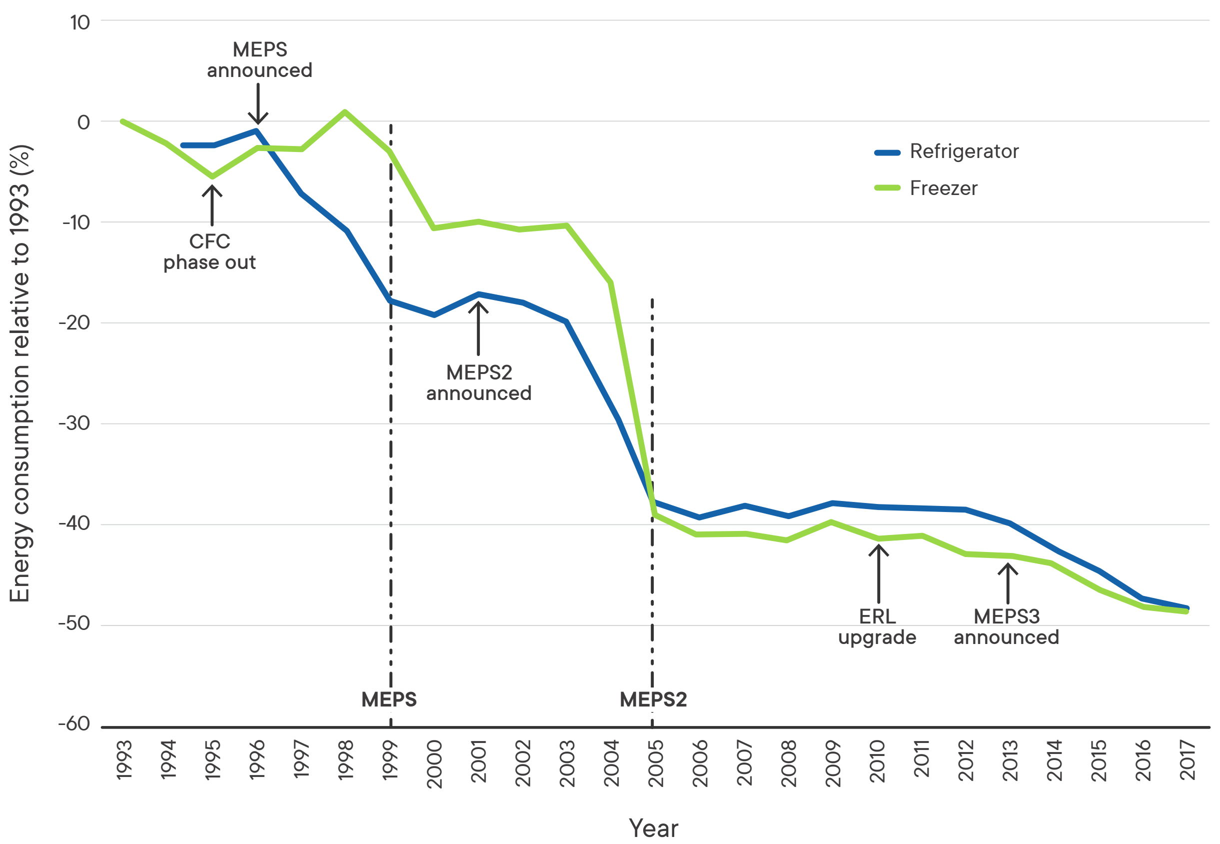
Reported outputs should include the national aggregated values for **net benefit** and the **benefit cost ratio**.

The former is key to giving the magnitude of the benefits, which can then be compared with those from other policies or for similar policies domestically or in other economies. The latter metric is a measure of overall cost effectiveness.

Reports should indicate whether a financial time series has been adjusted for inflation and, if so, what rate has been used. Real values, i.e. corrected for the value of inflation, should be used wherever possible. Also noted should be whether a discount rate has been applied and what rate has been used.

Figure 7: Improvements to refrigerator and freezer efficiency – Australia (1993-2017)

A good example of presenting the results of an evaluation is the Decision Regulation Impact Statement – Household Refrigerators and Freezers (E3 2017), which presents an analysis of the historical effect of MEPS and mandatory labels to date.



Source Figure 1 in E3 (2017).

KEY EVALUATION ISSUES:
Reporting

Other Outputs

Other outputs that may be requested include:

- national: cost benefits from improved health and including these in cost benefit calculations
- national: net effect on employment
- consumers: product choice (tracking the range of costs and features available on the market)
- consumers: average real cost of products over time (relating to affordability)
- suppliers: number of suppliers, changes of market share (indicative of profitability)
- power sector: changes in total energy and peak demand (feeds into energy, transmission & distribution costs and emissions)
- power sector: avoided marginal investment costs
- power sector: security of supply (how reduced energy demand can make supply more secure)
- local industry: market share
- local industry: range of products offered - in price and efficiency (both of these can feed into an assessment of competitiveness and therefore levels of revenue and employment).

Other very specific outputs may be requested, for example:

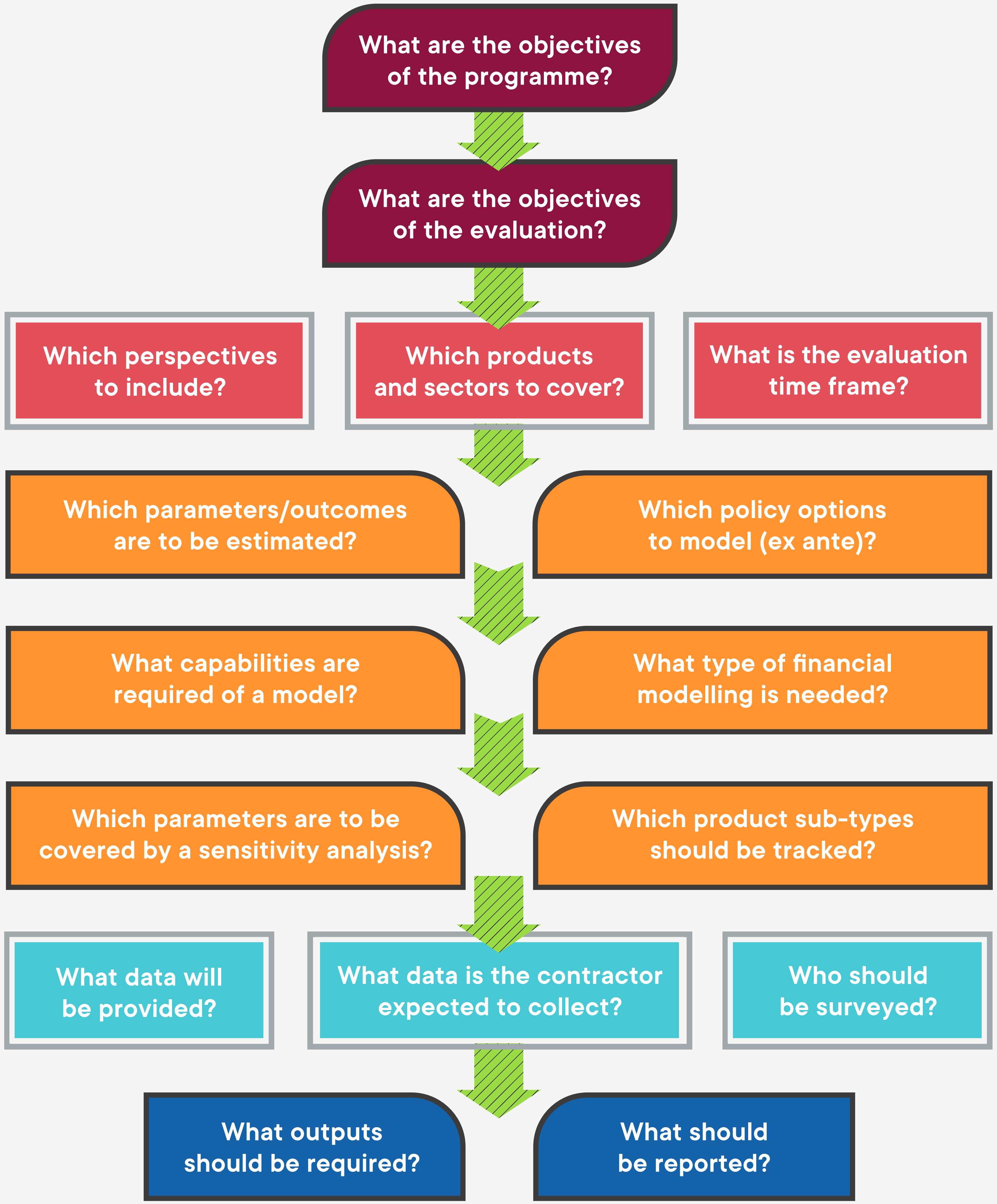
- the quantity of mercury used in fluorescent lighting to estimate the amount entering the waste stream, if hazardous waste is a focus
- the quantity and global warming potential (GWP) of refrigerants released during the lifetime of cooling equipment, if reducing GWP is a policy objective.

SUMMARY

Reporting

- Specify what inputs and outputs should be provided and in what form
- Consider providing a template to be completed
- Specify the content of the (public) report to include as much of the underlying data and assumptions as possible

Check list for commissioning a programme evaluation



Internal Resources

Commissioning an evaluation can be a resource intensive exercise, even where the bulk of work is undertaken externally. It is important that adequate management resources are allocated in addition to the direct cost of any external contractor or any internal work team.

The results of the evaluation will be used to set government policy in this area, so it is vital that the results are accurate, complete, robust and well presented. It will take sustained attention by the programme manager to effectively manage the overall evaluation project.

For example, it may be important to liaise with government departments other than those initiating the evaluation, for them to provide required input data, e.g.CO_2-e emissions, standard discount rates, and to account for their views. In these cases, it may be advisable for the programme manager to at least make introductions for the evaluators to their government colleagues or possibly undertake intra-governmental consultation directly. Another task for the programme manager may be to arrange access for the contractor to government held data or purchase of critical commercial data - this can also be time consuming.

Stakeholder consultation is a key part of many evaluations. This may form part of the contractors' responsibilities or may be undertaken separately by the lead government agency. In either case, the programme manager needs to be involved to add credibility and ensure all stakeholders are involved. This is why some governments prefer to manage this aspect of the process themselves, in which case resources, such as secretariat support and event management, will be needed.

All EES&L evaluations need significant input data. While all of this may be collected by the evaluator, if data needs to be purchased (for example market data from a commercial supplier) it may be preferable for this to be purchased by the programme manager, with associated direct and indirect (staff time) costs.

Most evaluations are public documents and the programme manager needs to allocate resources to present the results to government colleagues and to wider stakeholders - e.g. through press releases, on web pages, at meetings and webinars.

Finally it is crucial that enough time is allowed to plan, commission, undertake and present the evaluation results. Data collection and stakeholder consultation, in particular, can be a lengthy process. And while aspects can be accelerated, this may come with the risk of increased subcontractor costs, reduced quality or both. Allowing sufficient time, including allowing for some unexpected delays, will increase the chance of achieving a fully satisfactory evaluation.

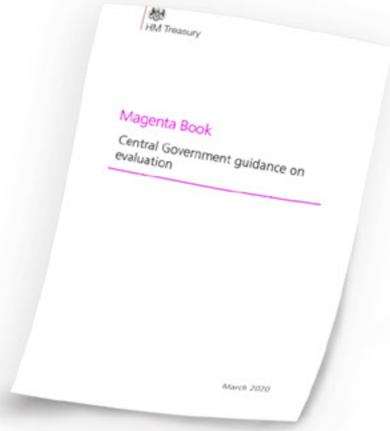
Other Resources

The following resources provide information on how to conduct formal evaluations as well as some specific guidance on specialised elements:

Energy-efficiency labels and standards: a guidebook for appliances, equipment, and lighting (CLASP 2005)

The Green Book: Central Government Guidance on Appraisal and Evaluation (UK HM Treasury 2022)

Protocols to conduct market and impact assessments (United for Efficiency (U4E) 2021b)



Magenta Book 2020 Central Government guidance on evaluation (various appendices and guides) (UK HM Treasury 2020)

Energy Labelling Guidance for Lighting and Appliances

Better Regulation Guidelines and **Better Regulation Toolbox** (Europe) (European Commission 2021a, 2021b)

Evaluation into Practice to Achieve Targets for Energy Efficiency (Europe) (EPATEE 2019)

Saving calculation methods and their application in the EPATEE Toolbox (Europe) (EPATEE 2022)

ANNEX A: Stock Models

This annex describes the details of stock models that are usually used to calculate impacts of an EES&L programme.

A CORE ELEMENTS OF A STOCK MODEL

In its simplest form, a stock model counts the number of pieces of new equipment entering the "stock" (i.e. the total pool of appliance or equipment in operation in the market) each year and adds these to the existing products already in operation from previous years. The normal time step for a stock model is one year - longer time steps are generally not suitable for appliances and equipment, which have lifetimes from 5 to 20+ years.

While new products are added to the stock each year, old products that have reached the end of their working life are removed each year from the pool of operating stock. Where energy attributes are improving over time, older products that use more energy leave the stock each year, while new products that use less energy enter the stock. If the stock numbers were stable, this would mean that overall stock energy would be decreasing. However, the stock model also has to be able to account for cases where the total stock is increasing or decreasing. This could be because the number of products per household or business is increasing (ownership) or because the total number of households or businesses is increasing (or combinations of both).

Importantly, a stock model should aggregate the number of new products entering the stock in each year multiplied by the energy attributes of the products entering in that year. This is added to the value calculated for the existing stock to give an overall stock-weighted average energy use in any particular year.

The stock model should also be able to take into account changes in usage over time. For example, if it is known that the hours of watching television is decreasing over time, then the decline in hours of viewing needs to be reflected in the model inputs and applied to the whole stock currently in service in any year (not just the new televisions entering the market) in order to get an accurate overall bottom-up estimate of total energy consumption. Other examples of usage changes are the mix of programme wash temperatures or average load size for clothes washers.

This means that the model energy attributes have to be constructed in a way that will allow usage factors to be externally applied to the whole stock when calculating total energy consumption. For a television, for example, the stock model would need to take account of on mode power and standby power. The hours of use can then be applied to the stock weighted power values to calculate the average Unit Energy Consumption (UEC) of the stock. Inclusion of features like automatic brightness control for TVs would be a more complex attribute to calculate.

For clothes washers, changes in wash temperature and loading level will have an impact on the per cycle energy consumption applicable to the entire stock.

B DEFINING TOTAL STOCK NUMBERS

Apart from product attributes and usage, a stock model has to track the total number of products in service over time. The most robust way of doing this is to calculate parameters like ownership or penetration over time and apply these to the total number of households or businesses each year.

Ownership and penetration are effectively the proportion of households or businesses with the product and are generally fairly stable and change slowly, so it is possible to project these forward with some confidence and to interpolate between historical survey data points, which may be spaced some years apart. The other advantage to this approach is that official data on the number of historical or future households or businesses can usually be obtained from official government sources, such as census data and government statistics, to provide more credible base data.

ANNEX A:
Stock Models

The key terms and definitions are:

- Penetration - the percentage of households or businesses that have one or more of a particular piece of equipment (maximum 100%)
- Ownership - the average number of pieces of equipment per household or business. This is effectively the average stock per household/ business (usually expressed as a decimal, can be more than 1.0)
- Saturation[19] - the average number of a particular piece of equipment that are installed and in use in homes/business that have the equipment installed (usually decimal, must be ≥ 1.0)
- Stock - this is a total count of a particular piece of equipment that is installed and in use.

Ownership is normally the most useful parameter for use in a stock model.

Penetration can be useful in some cases (e.g. number of households with a gas connection) when undertaking a detailed market analysis.

Useful equations are:

Stock = Ownership × number of households (or businesses)

Ownership = Saturation × Penetration

Where there is only one of a product in a business or household
(e.g. a dishwasher, washing machine, clothes dryers, hot water system)
then penetration and ownership will be the same and the saturation is 1.

It may be necessary to model some sub-types of appliances – for example, secondary televisions are usually smaller and older and used less than main televisions, so ownership data would need to split to the stock in primary and secondary appliances in order to get a more accurate picture of overall energy.

C PRODUCT LIFETIME AND RETIREMENT

The lifetime in a stock model is an estimate of the average expected life of products entering the stock. An overly simplistic stock model for a product lifetime of 10 years would retire all products entering the stock in 2010 in 2020.

However, most stock models assume some sort of distribution of retirements around an average value.

Simple estimates of average lifetime can be made if a sufficiently long series of stock and sales data are available.

Figure 8 illustrates three common approaches to retirement functions for a nominal product life of 15 years[20].

- A flat retirement assumes the same number of products are removed in each year of service after 1/3 lifetime up to 5/3 of lifetime.
- A linear retirement function draws a straight line from zero at 1/3 lifetime to a peak at the average lifetime then back to zero at 5/3 of lifetime.
- A normal distribution is a typical Gaussian function where the mean (average lifetime) and standard deviation[21] (spread) need to be defined.

19 Saturation in this context is different to the term market saturation, which reflects the degree of diffusion of a new or existing product into the market.
20 A Weibull distribution may also be used.
21 In a normal distribution, the standard deviation affects the spread of the distribution (larger standard deviation makes the bell curve wider and flatter). Experience has shown that a standard deviation of around one quarter of the expected life gives a reasonable distribution for many types of appliances and equipment.

ANNEX A:
Stock Models

In practical terms, these all look fairly similar: all units installed in year 0 would be expected to remain in service for at least 6 years and then units would be progressively removed from the stock as time progresses, with virtually no units remaining in service after 25 years (see Figure 9).

Most appliances and equipment will be expected to have a nominal lifetime defined in terms of years. However, there will be some products where the lifetime will be a function of usage. The most notable example is for lighting, where lifetime is usually defined in terms of operating hours. In this case, a stock model may have to take into account hours of use over time and include ways of varying the calculated elapsed lifetime as hours of use increase or decrease in different parts of the house or office.

Another product that is likely to have a lifetime that is at least partially impacted by hours of use is electric motors. Product experts would need to be consulted to see whether this needs to be factored into the relevant stock model.

For most products, it is assumed that the product lifetime will remain constant over the evaluation timeframe.

Average lifetime can also vary by product sub-type, for example, separate freezers may have a longer average lifetime than refrigerator-freezers. Secondary televisions may have a shorter effective lifetime.

However, there may be cases where the expected lifetime varies with the year of installation, if products are of poorer quality over time, or vice versa. This may especially be the case where policy measures also affect product lifetime.

While it is mathematically possible to apply a varying product lifetime to different cohorts, this is more complex than applying a fixed lifetime and may require development of more specialised software. Thus, there would need to be compelling evidence of lifetime varying by year of product sale before embarking on this approach.

Figure 8: Three common retirement functions for product lifetime

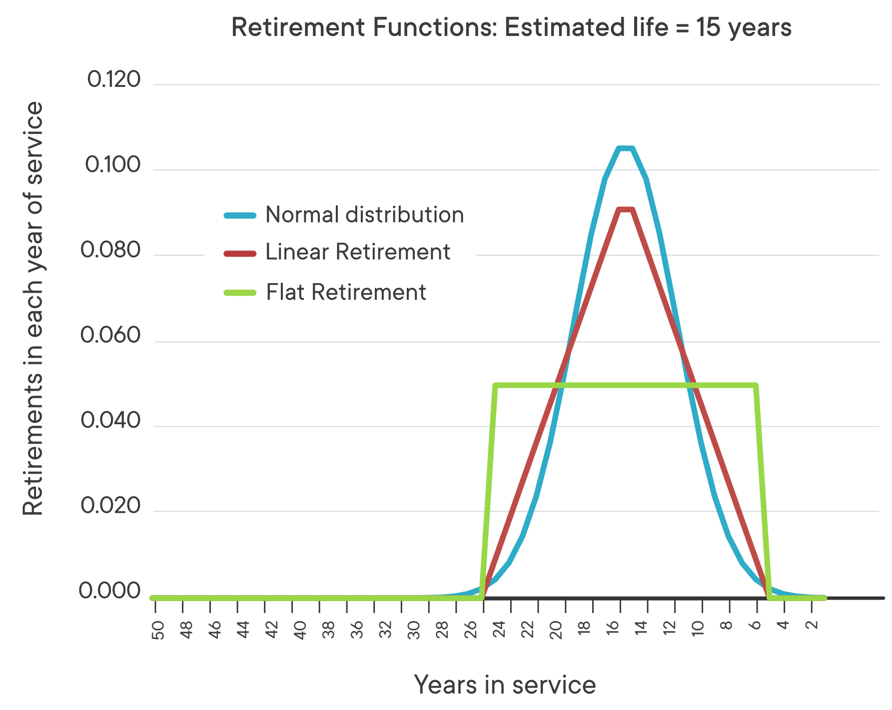
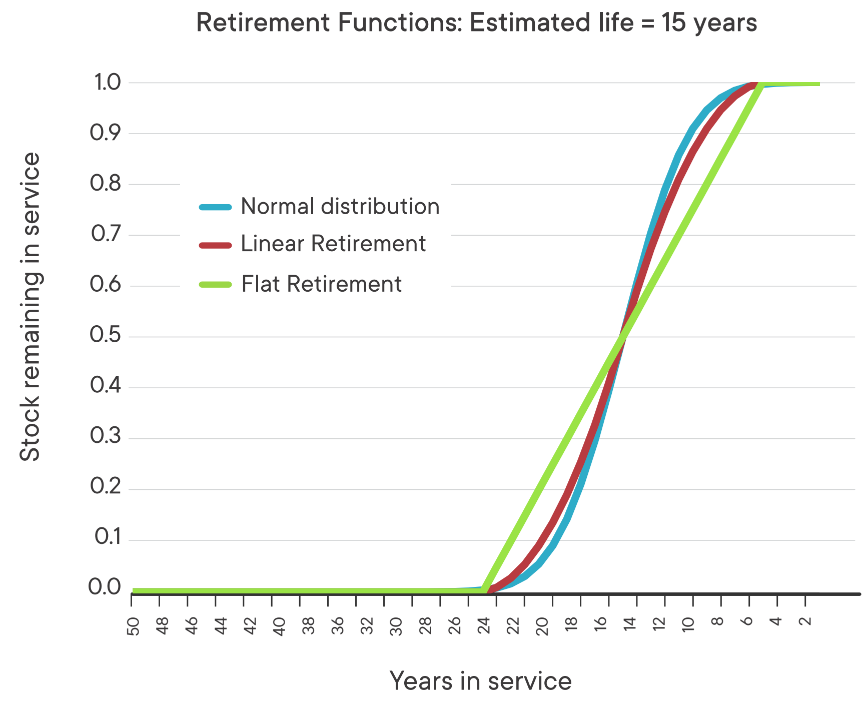


Figure 9: Estimating stock remaining in service



ANNEX A:
Stock Models

D RECONCILING SALES, STOCK CHANGES AND PRODUCT LIFETIMES

A stock model is intended to provide a mathematical representation of what happens in real life, so the stock model should be calibrated against known datasets wherever possible.

In simple terms, new products sold are entering the stock each year. Some of these are replacing existing stock that has reached the end of its working life and some are new products in households or businesses, due to increasing ownership and/or increasing numbers of households or businesses.

As the product lifetime is increased, the number of products being replaced each year will decrease and vice versa. The number of products being installed in new-only installations are not affected by changes in lifetime. If there is good data on product sales, this provides an opportunity for the stock model to be calibrated[22] by comparing the sales generated by the stock model and the actual known sales.

Sales data is more difficult to use in this way if products are being sold into different sectors and the sales data is not split by sector e.g. air conditioners, televisions and computers, which typically have substantial sales into both the household and business sectors.

Typically, the actual stock in a stock model should be anchored to the externally supplied data series (based on household and business numbers times ownership). Lifetime values should be based on the best available external data, but with a cross check on the sales stream generated by the stock model where this can be compared to actual data. Real sales vary from year to year due to changes in economic conditions[23], whereas the sales generated by a stock model tend to be more constant as these macroeconomic conditions are not reflected in this type of model.

The sales stream generated by the stock model is used to weight the product attributes being tracked.

E MODELLING PRODUCT SUB-TYPES

A stock model for a specific end use product is often split into a number of sub-models that reflect characteristics of particular product types or different usage parameters. When splitting into sub-models, it is important to ensure that all of the sub-models sum to provide an accurate value for the overall product stock.

The use of sub-models within a stock model are typically required where the energy attributes, lifetime, energy efficiency and/or usage of different product sub-types is significantly different. Usually, different sectors will require separate sub-models.

For example, a clothes dryer stock model would sensibly be split into heat pump dryers and resistance element dryers, as their energy characteristics are very different and the mix of sales by type may be changing over time. Also, heat pump dryers may be self-selected more by heavy users due to their lower operating costs.

Similarly, for clothes washers, there may be sub-models for top loading (vertical axis) and front loading (horizontal axis) clothes washers as their energy and water characteristics are very different.

22 If any two of sales, ownership and lifetime are known with complete data in a time series, then it should be possible to derive the other parameter by calibrating the stock model. However, time series that are long and complete enough are rarely available and there are other factors (such as economic conditions) that will affect sales in certain years - these cannot be reflected in this type of bottom up model.
23 A recession can reduce sales of both essential and non-essential equipment types. For example, it is known that periodical events such as the Olympics can have a marginal cyclic impact on television sales.

ANNEX A:
Stock Models

Some suggested stock sub-model splits for different appliance and equipment types include:

- **heating:** gas room, gas central, air conditioner room (split and/or window wall), air conditioner central, wood, oil, LPG, district heating. These may also need to be split into climate zones
- **cooling:** room air conditioner (split and/or window wall), central air conditioner, room evaporative, central evaporative. These may also need to be separated into climate zones
- **water heaters:** gas storage, gas instantaneous, electric storage (small and/or large, restricted energisation), electric instantaneous, solar thermal electric boost (flat plate or evacuated tube), solar thermal gas boost, heat pump, solid fuel (wood), district heating (also combi space and water heating systems). Ambient systems (solar thermal and heat pump) may need to be split into climate zones
- **clothes washers:** top loading (vertical axis), front loading (horizontal axis)
- **clothes dryers:** heat pump, resistance electric (may consider a split on auto sensing versus timer control)
- **computers and monitors:** laptops, desktops; residential, commercial and professional (usually very different usage profiles)
- **televisions:** residential, commercial applications (usually very different usage profiles), possible primary and secondary televisions
- **household refrigeration:** refrigerator-freezers, refrigerators, separate freezers. These may also need to be split into climate zones
- **commercial refrigeration:** self-contained, remote; sold door, glass door, open; chilled, frozen; cool rooms; professional refrigeration, retail display
- **lighting:** LED, fluorescent, High Intensity Discharge (HID); commercial, residential, outdoor, road. Many of these have different lifetimes so this would also necessitate the use of separate sub-models for each technology.
- **distribution transformers:** voltage, phase, kVA capacity (loading will vary by size and type); dry, liquid filled
- **electric motors:** phase, capacity, number of poles (average usage, typically varies a lot by capacity).

F AGGREGATING MODEL OUTPUTS TO CALCULATE STOCK WEIGHTED ENERGY AND ATTRIBUTES

Once the stock model (and all of the sub-models) are calibrated and are working consistently, the stock model can be used to aggregate results to a regional or national level.

Each model or sub-model should produce data on weighted average energy characteristics and other important attributes such as size for each year of the evaluation timeframe (typically 25 to 50 years).

The stock model then applies any user and/or climate related parameters to the stock each year to calculate a weighted UEC (by fuel where applicable) in each year. This is then scaled up by the stock installed each year to calculate total energy consumption for the sub-model or model.

Aggregated energy data can then be used to calculate energy related emissions (using the relevant emission intensity factors for each year), total new product purchase costs, total energy costs and any other relevant parameters.

A stock model typically generates all of these outputs separately for each case that is examined (e.g. “without measures", “with measures” cases). The key aggregated outputs of energy, emissions, product costs and energy costs can then be collated for further high level analysis.

The start year and the end year should be carefully specified. The start year is the year where all scenarios (“without measures” and “with measures” cases) all have the same aggregated data (i.e. the difference between scenarios is zero). The end year is the last year that products are assumed to be installed in the stock model.

ANNEX A:
Stock Models

For example, we may wish to compare two scenarios where products are installed up to and including 2035. The stock model would generate annual outputs for all key parameters to 2035. Given that new products are installed in 2035 and the stock model is counting the purchase cost of products installed in 2035, it would be normal to specify that the energy consumption profile of products installed in 2035 be allowed to run until all of these are retired from the stock (i.e. until the remaining energy consumption is zero).

In this example, the stock model would need to continue calculating energy for an additional 25 years (to 2060) if the product average life was 15 years. If the model data was just truncated at 2035, almost none of the energy benefits from more efficient products installed from around 2030 would be counted but all of the purchase costs would be counted. This would give a biased assessment (on the assumption that more efficient products have a higher purchase cost and a lower operating cost).

The following tables show some sample stock model data for a particular year (2022). The key inputs are:

- stock remaining percentage (this is the proportion of new products installed in previous years that are still operating in the stock): this is calculated from the selected retirement function
- sales in the specific year
- stock in service (in the current year 2022): this is calculated from sales in year × stock remaining percentage in that year
- Unit Energy Consumption (UEC): this is the expected energy consumption of the products installed in a particular year, adjusted for any changes in external conditions and usage factors over time and in response to policy changes
- total energy consumed = number in the stock × UEC.

This type of calculation can also determine stock weighted characteristics such as size or capacity, which is useful to track over time. The average capacity for each cohort is multiplied by the remaining stock in service in that year and summed for all years. This value is then divided by the total stock in service to provide a stock weighted average capacity.

While the stock model shown is a relatively simple set of calculations, this includes the UEC profile for just one set of policies ("with measures" in this case). This set of calculations also only gives the total energy consumption in the modelling end year shown (2022 in this case). These calculations would have to be replicated for each year being modelled (say 2010 to 2060) in 1-year steps and for each set of policy scenarios ("without measures" and various "with measures" cases). This is for just a single product, so if there were multiple sub-types being modelled, these would all have to be run in parallel. A full stock model may have to run many hundreds of stock calculation sets to get the full range of energy outputs over the required time frame.

Running in parallel with these calculations are the financial calculations that determine energy costs and emissions in each year, total product purchase costs under different policy scenarios and any other financial calculations (external costs of emissions, cost streams under different discount rates and so on).

ANNEX B:
Sample Input and Output Tables

These examples are taken from the UK's Market Transformation Programme Briefing Notes. They are examples of ex-ante evaluations, which were published in 2009.

A EXAMPLE FOR DOMESTIC COLD APPLIANCES

Input data

Table 1: Sample summary of stock figures for cold appliances

Year	Chest freezers	Upright freezers	Fridges	Fridge-freezers
2008	4,193,310	7,818,810	9,720,603	17,864,892
2010	4,180,986	8,115,400	9,913,965	18,220,260
2020	4,418,527	9,132,091	10,899,497	20,031,508
2030	4,802,991	9,926,690	11,847,881	21,774,484

Table 2: Sample cold appliance total sales

Year	Chest freezers	Upright freezers	Fridges	Fridge-freezers
2008	253,652	542,025	833,995	1,244,120
2010	272,499	578,715	798,481	1,239,018
2020	281,369	620,575	897,812	1,396,940
2030	313,988	685,038	957,431	1,485,911

Table 3: Sample average predicted sales-weighted energy consumption for new cold appliances – base scenario

Average efficiency new appliances (kWh/year)	2007	2010	2020	2030
Chest freezers	296	176	141	141
Upright freezers	253	226	168	168
Fridges	175	158	121	121
Fridge-freezers	354	327	261	261

Table 4: Sample average predicted sales-weighted energy efficiency index for new cold appliances – policy scenario

Average efficiency new appliances	2009	2020	2030
Chest freezers	71.2	31.5	28.6
Upright freezers	56.6	31.8	28.9
Fridges	52.4	30.5	26.0
Fridge-freezers	56.2	31.4	29.5

ANNEX B:
Sample Input and Output Tables

Output data

Table 5: Sample summary base case outputs

Energy Consumption (GWh)	2009	2020	2030
Upright freezers	2620	1880	1690
Chest freezers	1450	850	690
Fridge freezers	8370	6230	5750
Refrigerators	2050	1510	1430
TOTAL	**14490**	**10470**	**9570**
CO_2 Emissions ($MtCO_2$)			
Upright freezers	0.95	0.69	0.62
Chest freezers	0.53	0.31	0.26
Fridge freezers	3.04	2.30	2.12
Refrigerators	0.75	0.56	0.53
TOTAL	**5.27**	**3.86**	**3.53**

Table 6: Sample summary policy scenario outputs

Energy Consumption (GWh)	2009	2020	2030
Upright freezers	2620	1730	1330
Chest freezers	1450	810	550
Fridge freezers	8370	5740	4320
Refrigerators	2050	1350	1110
TOTAL	**14480**	**9630**	**7310**
Energy Savings (GWh)			
Upright freezers	0	150	360
Chest freezers	0	40	140
Fridge freezers	0	490	1430
Refrigerators	0	160	320
TOTAL	**0**	**840**	**2260**
CO_2 Emissions ($MtCO_2$)			
Upright freezers	0.95	0.64	0.49
Chest freezers	0.53	0.30	0.20
Fridge freezers	3.04	2.12	1.60
Refrigerators	0.75	0.50	0.41
TOTAL	**5.27**	**3.55**	**2.70**
CO_2 Emissions Savings ($MtCO_2$)			
Upright freezers	0.000	0.06	0.13
Chest freezers	0.000	0.02	0.05
Fridge freezers	0.002	0.18	0.53
Refrigerators	0.000	0.06	0.12
TOTAL	**0.002**	**0.31**	**0.83**

ANNEX B:

Sample Input and Output Tables

B EXAMPLE OF PACKAGED AIR CONDITIONERS

Input data

Table 7: Sample summary Packaged Air Conditioning Unit Stock Data

Stock – All Scenarios								
	TOTAL	Close Control	Ducted Split	Indoor	Mini Split	Moveable	Roof Top	Window
2008	2,136,000	84,000	3,600	50,000	1,482,000	459,000	25,000	33,000
2010	2,120,000	87,000	3,700	55,000	1,502,000	421,000	25,000	28,000
2020	3,860,000	91,000	3,600	60,000	3,246,000	421,000	28,000	11,000
2030	5,105,000	96,000	3,700	70,000	4,410,000	467,000	52,000	5,200

Table 8: Sample summary Packaged Air Conditioning Sales

Sales – All Scenarios								
	TOTAL	Close Control	Ducted Split	Indoor	Mini Split	Moveable	Roof Top	Window
2008	238,000	9,400	280	4,700	168,000	52,000	2,000	1,600
2010	237,000	7,100	270	3,800	172,000	51,000	1,500	1,500
2020	548,000	7,400	280	5,200	464,000	67,000	3,300	1,000
2030	649,000	8,100	310	6,400	554,000	74,000	6,000	370

ANNEX B:
Sample Input and Output Tables

Output data

Table 9: Sample summary of energy consumption and savings and CO_2 emissions and savings

Energy Consumption policy scenario (GWh)	2009	2020	2030
Window	30	10	0
Close	2130	1740	1520
Ducted-split	20	10	10
Indoor	20	20	10
Mini-split	3040	5010	6370
Moveable	180	100	110
Roof-Top	510	430	680
TOTAL	**5930**	**7310**	**8700**
Energy Savings (GWh) (difference between base case and policy scenarios)	**2023**	**2030**	**2040**
Window	0	0	0
Close	0	530	870
Ducted-split	0	0	0
Indoor	0	0	0
Mini-split	0	1930	3100
Moveable	0	30	30
Roof-Top	0	80	170
TOTAL	**0**	**2580**	**4170**
CO_2 Emissions ($MtCO_2$)	**2009**	**2020**	**2030**
Window	0.01	0.00	0.00
Close	0.92	0.75	0.66
Ducted-split	0.01	0.00	0.00
Indoor	0.01	0.01	0.01
Mini-split	1.31	2.16	2.74
Moveable	0.08	0.04	0.05
Roof-Top	0.22	0.18	0.29
TOTAL	**2.55**	**3.14**	**3.74**
CO_2 Emissions Savings ($MtCO_2$) (difference between base case and policy scenarios)	**2009**	**2020**	**2030**
Window	0.00	0.00	0.00
Close	0.00	0.23	0.37
Ducted-split	0.00	0.00	0.00
Indoor	0.00	0.00	0.00
Mini-split	0.00	0.83	1.33
Moveable	0.00	0.01	0.01
Roof-Top	0.00	0.04	0.07
TOTAL	**0.00**	**1.11**	**1.79**

References

CLASP 2005, Standards & Labelling Guidebook 2nd ed, CLASP, Washington, DC, retrieved from https://clasp.ngo/tools/s-l-guidebook.

Desroches, L-BG 2013, 'Incorporating experience curves in appliance standards analysis', Energy Policy, vol. 52, pp. 402–416.

E3 2017, Decision Regulation Impact Statement – Household Refrigerators and Freezers, Department of the Environment and Energy, Canberra, Australia, retrieved from http://www.energyrating.gov.au/news/decision-regulation-impact-statement-household-refrigerators-and-freezers.

Ellis, M, Jollands, N, Harrington, L & Meier, AK 2007, 'Do energy efficient appliances cost more?', in ECEEE Summer Study on Energy Efficiency, Mandelieu, France, retrieved from https://www.eceee.org/static/media/uploads/site-2/library/conference_proceedings/eceee_Summer_Studies/2007/Panel_6/6.025/paper.pdf.

Energy Efficient Strategies 2016, Whitegoods Efficiency Trends in Australia 1993-2014, Department of Innovation, Industry and Science, Canberra, Australia, retrieved from http://www.energyrating.gov.au/document/whitegoods-efficiency-trends-1993-2014.

Energy Efficient Strategies, Waide Strategic Efficiency, Ballarat Consulting & CLASP 2021, Achievements of Energy Efficiency Appliance and Equipment Standards and Labelling Programs: 2021 update, IEA 4E, retrieved from https://www.iea-4e.org/wp-content/uploads/2021/09/EESL2021_Main-Report.pdf.

EPATEE 2019, Evaluating legislative/normative measures on energy efficiency of Household appliances using stock modelling, Evaluation into Practice to Achieve Targets for Energy Efficiency, retrieved from https://www.epatee-toolbox.eu/wp-content/uploads/2019/10/EpateeGuide-05.pdf.

EPATEE 2022, 'Saving calculation methods and their application in the EPATEE Toolbox', P, retrieved from https://epatee.eu/online-tool-guidance-and-support-put-evaluation-energy-savings-programs-practice.

European Commission 2021a, Better Regulation Guidelines, Brussels, retrieved from https://ec.europa.eu/info/law/law-making-process/planning-and-proposing-law/better-regulation-why-and-how/better-regulation-guidelines-and-toolbox_en.

European Commission 2021b, 'Better Regulation Toolbox', retrieved from https://ec.europa.eu/info/law/law-making-process/planning-and-proposing-law/better-regulation-why-and-how/better-regulation-guidelines-and-toolbox_en.

Fischer, A 2021, 'How Energy Efficiency Will Power Net Zero Climate Goals', retrieved from https://www.iea.org/commentaries/how-energy-efficiency-will-power-net-zero-climate-goals.

Harrington, L 2017, 'Tracking the energy efficiency of whitegoods in Australia', in International Conference on Energy Efficiency in Domestic Appliances and Lighting, European Commission, University of California, Irvine, USA.

Harrington, L, Lane, K 2010, Evaluation of Energy Efficiency Policy Measures for Household Refrigeration in Australia, prepared for the Department of Climate Change and Energy Efficiency, December 2010.

International Energy Agency 2021, Energy Efficiency 2021, Paris, France, retrieved from https://www.iea.org/reports/energy-efficiency-2021.

References

IEA/4E TCP 2021, Achievements of Energy Efficiency Appliance and Equipment Standards and Labelling Programmes, IEA, Paris (2021), License: Creative Commons Attribution CC BY-SA 3.0 IGO https://www.iea.org/reports/achievements-of-energy-efficiency-appliance-and-equipment-standards-and-labelling-programmes.

Mogensen, KS, Blomquist, L, Bennich, P, Stengård, L & Fjordbak, T 2019, 'NORDCRAWL – The Nordic framework for collection, analysis and surveillance of market data based on automated and frequent crawling of retail web shops', in ECEEE, France, retrieved from https://www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2019/9-improving-energy-efficiency-in-ict-appliances-and-products/nordcrawl-the-nordic-framework-for-collection-analysis-and-surveillance-of-market-data-based-on-automated-and-frequent-crawling-of-retail-web-shops/.

Motherway, B 2019, 'Energy efficiency is the first fuel, and demand for it needs to grow', retrieved from https://www.iea.org/commentaries/energy-efficiency-is-the-first-fuel-and-demand-for-it-needs-to-grow.

UK HM Treasury 2020, Magenta Book 2020 Central Government guidance on evaluation, UK Government, London, United Kingdom, retrieved from https://www.gov.uk/government/publications/the-magenta-book.

UK HM Treasury 2022, The Green Book: appraisal and evaluation in central government, UK Government, London, United Kingdom, retrieved from https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government.

United for Efficiency (U4E) 2021, Energy Labelling Guidance for Lighting and Appliances, retrieved from https://united4efficiency.org/wp-content/uploads/2021/01/U4E-Labelling-Guidance_20210125.pdf.

United for Efficiency (U4E) 2021b, Protocols to Conduct Market and Impact Assessments, retrieved from https://united4efficiency.org/resources/protocols-to-conduct-market-and-impact-assessments/.

US Department of Energy 2011, Using the Experience Curve Approach for Appliance Price Forecasting.

Waide, P 2000, COLD II The revision of energy labelling and minimum energy efficiency standards for domestic refrigeration appliances - Final report, ADEME, France.

Waide, P 2022, 'Personal communication'.

Weiss, M, Patel, MK, Junginger, M & Blok, K 2010, 'Analyzing price and efficiency dynamics of large appliances with the experience curve approach', Energy Policy, vol. 38, pp. 770–783.