Chapter 8 – Methodologies for analysing impacts in impact assessments, evaluations, and fitness checks

Chapter 8 presents a **non-exhaustive list of relevant methodologies for policy analysis** in impact assessments and evaluations [812]. According to the 'better regulation' guidelines, significant impacts should be assessed quantitatively if possible as well as qualitatively. The analytical methods in Chapter 8 cover a wide spectrum of approaches and contribute to ensure effectiveness, consistency, and transparency of the policy process.

Every **impact assessment** aims to provide answers to two fundamental questions (see questions 5 and 6 in the 'better regulation' guidelines Chapter IV), namely:

- What are the impacts of the different policy options and who will be affected?
- How do the options compare?

Correspondingly, **evaluations and fitness checks** (see question 3 in the 'better regulation' guidelines Chapter III) aim to reply to:

- To what extent was the intervention successful?

The methods comprise both forward-looking and backward-looking approaches, although several tools are relevant for both. Furthermore, ensuring coherence across the policy cycle implies considering the links that connect the various phases. If the intervention logic and the planning of data are well done in the ex-ante phase, the later evaluation will be easier to conduct. Conversely, a good evaluation can and should feed into an impact assessment.

The choice of which impacts to analyse and methods to use should be guided by the principle of proportionate analysis (see Tool #12 (*How to apply proportionality to impact assessments*)). This means that the scope and depth of the analysis should be proportionate and consistent with the importance and type of initiative as well as with the nature and magnitude of the expected impacts. The effort invested in data collection, the depth of analysis, the extent of quantification and thereby choice of analytical method(s) should correspond to the likely magnitude of impacts.

A preliminary screening of potential impacts across the social, economic, and environmental dimensions should identify those that are likely to be most significant or critical for political decision-making for each specific initiative.

Guiding questions outlining the criteria for choosing the most appropriate analytical approach

1. **What are the most significant effects and impacts? Which method can be applied?**
   Not every method is equally suited to analyse the specific effects or impacts of a measure. Often a single measure needs various (inter-linked) methods (e.g. to analyse economic, social and environmental impacts)
2. **Which type of problem structuring is required?**
   Which are the steps to implement the method for a real-world measure?
3. **What is the time horizon of the expect effects of the intervention?**
   The method should cover the full time horizon with significant effects.

[812] See Tool #51 for an overview of methods to consult stakeholders, both in open and targeted manner – such as interviews, focus groups, seminars, Eurobarometer surveys and others.

4. **What kind of data/information (quantitative, qualitative …) are required to analyse the problem?**
   Is the required data available and accessible? How sensitive is the method for (partially) missing data?
5. **How big is the amount of information/data required by the method? Which technical tools (software, databases etc.) are necessary to apply the method?**
   Are the necessary technical tools, knowledge, and hardware available to handle the required amount of data? If not, can they be acquired?
6. **What kind of knowledge/expertise is required on the part of the analyst?**
   Is the required knowledge/expertise available and if not, can it be acquired?
7. **How many resources does the method need?**
   Is the investment in a method proportional to the size of the problem? Can the results or methods be used (again) for other (future) problems (including in other DG's)?
8. **How long does it take to apply the method?**
   What is the estimated timeframe of the analysis and does it correspond with the objectives or deadlines of the political process? What are the strategies if there is a lack of time?
9. **Has the method been used in previous IAs/evaluations/fitness checks/studies?**
   Who might have experience in using the method and can provide advice?

The **Joint Research Centre** provides operational guidance and assistance on methodologies.

- The **Commission Competence Centre on Modelling** (CC-MOD) provides advice on model use for policymaking; its Modelling Inventory and Knowledge Management System (MIDAS); uncertainty and sensitivity; transparency and policy coherence with baseline scenarios; social multi-criteria evaluation of policy options; and peer-reviews by external experts.
  Contacts: EC-CCMOD@ec.europa.eu
  https://ec.europa.eu/knowledge4policy/modelling_en
- The JRC supports DGs with their **Life Cycle Assessments** studies at product and meso/macro scale (Tool #66). **Contacts:** JRC-ENVIRONMENTAL-FOOTPRINT@ec.europa.eu
- The Commission **Competence Centre on Microeconomic Evaluation** (CC-ME) provides advice and support for ex-post evaluation for policymaking (Tools #67, #68).
  Contacts: EC-CC-ME@ec.europa.eu
  https://ec.europa.eu/knowledge4policy/microeconomic-evaluation_en
  https://microeconomicevaluation.jrc.ec.europa.eu/ (tools)
- The Commission **Competence Centre on Foresight** (CC-FUTUR) provides advice on the use of foresight and knowledge bases, such as EC Megatrends Hub, Horizon Scanning (Tool #20).
  Contacts: JRC-FORESIGHT@ec.europa.eu
  EU Policy Lab: https://blogs.ec.europa.eu/eupolicylab/
  EC Megatrends Hub: https://ec.europa.eu/knowledge4policy/foresight_en
  Megatrends Workshop: https://ec.europa.eu/knowledge4policy/foresight/megatrends-engagement-tools_en
  JRC Horizon Scanning Network: https://webgate.ec.europa.eu/connected/groups/horizon-

scanning

- The Commission **Competence Centre on Behavioural Insights** (CC-BI) combines insights from different disciplines to support EU policymaking by identifying behavioural elements in policies and proposing behavioural levers to increase their effectiveness. (Tool #69).
  Contacts: JRC-CCBI@ec.europa.eu
  https://ec.europa.eu/jrc/en/research/crosscutting-activities/behavioural-insights
- The Commission **Competence Centre on Text Mining and Analysis** (CC-TMA) uses text mining and analysis tools to extract information from online data, including traditional or social media, or from large public or proprietary document sets.
  Contacts: JRC-TMA-CC@ec.europa.eu
  https://knowledge4policy.ec.europa.eu/text-mining/about_en

TOOL #56. TYPOLOGY OF COSTS AND BENEFITS

1. INTRODUCTION

To assess the impacts of a policy option − both ex ante in impact assessments and ex post in evaluations / fitness checks − one needs to look to as many benefits and costs as possible, in a proportionate manner. Depending on the nature of the impacts, these costs and benefits can be direct or indirect. This tool describes the different types of costs and benefits, and Table 1 lists them[813].

Costs related to an initiative/intervention, often are concentrated on a specific group of stakeholders. Benefits, on the other hand, tend to emerge over a longer period and may be more society-wide.

At times, what is a cost to one party may be a benefit to another and these symmetrical changes in private welfare will normally cancel out at the societal level. In addition, investments, which are needed to comply with legislation, generate at the same time economic activity and income while (possibly) enabling cost savings later. It is therefore very important to distinguish between costs and benefits that represent net additions or reductions of total welfare for the society, as opposed to costs and benefits that arise for specific categories of stakeholders as a result of a transfer of resources. It is equally important to report both on the net societal effect and the distributional effects on individual groups in the society as this may determine the type of measures considered in the initiative.

Costs and benefits can be aggregated and compared more easily when expressed in the same units. Economists tend to favour monetary terms, using 'market prices'. However, many (equally important) impacts are not easily expressed in monetary terms (e.g. what is the monetary value of fundamental rights or biodiversity?). In some cases, methods can help to convert these into impacts in monetary terms (see Tool #57 (*Methods to assess costs and benefits*)).

2. CATEGORIES OF COSTS AND BENEFITS

For the application in the 'better regulation' policy[814], three main categories of costs and benefits can be distinguished:

- **Direct** and **indirect**
  Direct costs/benefits can be clearly linked to the intervention as they result from the respective legal provisions.

  Indirect costs/benefits are observed in related upstream or downstream markets or experienced by consumers, government or other stakeholders (e.g. third countries) that are not directly targeted by the initiative/regulation.
- **Private** and **societal**
  Costs and benefits can be analysed from the perspective of citizens, consumers, businesses, or public administrations. Impact assessments look at all of these together

[813] Assessing the Costs and Benefits of Regulation; CEPS (2013) a study prepared for the European Commission; Chapter 1.

[814] In other domains, the considered categories of costs will differ depending on the purpose. For instance, in finances or accounting one can distinguish between e.g. fixed, variable, operational, marginal, sunk, and (un)controllable costs.

to provide an analysis from the point of view of society as a whole, whilst identifying the winners and losers.

- **One-off** and **recurrent**
  Depending on the point in time in which costs have to be borne or benefits occur, all the costs and benefits can be either 'one-off' or 'recurrent'.
  One-off costs/benefits often appear shortly after an initiative is implemented (including the grace periods where relevant).
  Recurrent costs/benefits appear more or less regularly over a certain period of time (or indefinitely) and they should be discounted to take into account the time preference (see Tool #64 (*Discount factors*)).

Figure 1. A map of regulatory costs and benefits

	Costs		**Benefits**	
Direct	Direct compliance costs	Adjustment costs Administrative costs Charges	Improved welfare	Health Safety Environment Direct economic benefits
	Enforcement costs	Information and monitoring Complaint handling Inspections Adjudication/litigation	Improved market efficiency	Cost savings Improved information Wider range of products/services
	Hassle costs			
Indirect	Indirect compliance costs		Indirect compliance benefits	
	Other indirect costs	Substitution effects Transaction costs Opportunity costs Negative effects on market functioning	Wider economic benefits	
			Other, non-monetary benefits	
Aggregate impacts				

3. DIFFERENT TYPES OF COSTS

Next to observing the occurrence of costs in time, the most useful distinction from the regulatory perspective is to look at direct and indirect costs of a given initiative; taken together, they represent the total cost of intervention:

(1) Direct costs of regulation

- *Direct compliance costs* are the costs that need to be borne to comply with the provisions of the regulation. They are composed of:
  - *Adjustment costs*, which encompass those investments and expenses that businesses, citizens, or public authorities have to bear in order to adjust their activity to the requirements contained in a legal rule;
  - *Administrative costs* are those costs borne by businesses, citizens, civil society organisations and public authorities as a result of administrative activities performed to comply with administrative obligations included in legal rules;
  - *Regulatory charges*, which include fees, levies, taxes, etc.
- *Enforcement costs* are associated with activities linked to the implementation of an initiative such as monitoring, inspections and adjudication/litigation.
- *Hassle costs* are often interpreted as 'regulatory annoyance' resulting from unnecessary waiting time, delays, redundant legal provisions, corruption, etc. They can apply equally to businesses, consumers, and citizens. As this category of costs is not well-defined, in most cases it is not analysed in impact assessments, evaluations and fitness checks.

(2) Indirect costs of regulation – incurred in related markets or experienced by consumers, government agencies or other stakeholders that are not directly targeted by the initiative/regulation.

Indirect costs are usually transmitted through changes in the prices and/or availability and /or quality of the goods or services produced in the regulated sector. Changes in these prices then ripple through the rest of the economy changing prices in other sectors and ultimately affecting the purchasing power of consumers and activities of economic sectors. The category includes:

- *indirect compliance costs* – cost related to the fact that other stakeholders have to comply with legislation;
- *offsetting/substitution costs* – for example costs related to reliance on alternative sources of supply;
- *transaction costs* – costs of renegotiating contracts;
- *opportunity costs* – the costs of foregone alternative investments; and
- *negative impacts on market functioning* such as reduced competition or market access, reduced innovation or investment.

Different policy instruments (see Tool #17) usually imply the existence (and therefore analysis) of different cost categories (Table 2).

Table 1: Policy types and associated recurrent costs

Type of regulatory alternative	Recurrent costs
Self-regulation	Monitoring costs Transaction costs Direct compliance cost
Co-regulation	Monitoring costs Enforcement costs Transaction costs Direct compliance cost
Market-based instruments	Transaction costs Charges Direct compliance costs Indirect compliance costs
Performance-based standards	Monitoring costs Direct compliance costs Indirect compliance costs
Command and control	Charges Administrative burdens Direct compliance costs Indirect compliance costs Monitoring costs Enforcement costs Adjudication

4. IMPACT OF REGULATORY COSTS ON DIFFERENT STAKEHOLDERS

Table 2 categorise costs in terms of major stakeholders i.e. business, citizens and consumers and public administrations. For individual initiatives this classification may vary somewhat.

Table 2: Impact of regulatory costs on different stakeholders

Type of cost		Citizens	Consumers	Business	Administrations
Direct costs					
Direct compliance costs	Adjustment costs	•		•	•
	Administrative costs	•		•	•
	Charges		•	•	
Enforcement costs	Information & monitoring	•		•	•
	Inspections and sanctions			•	•

	Complaint handling	•		•	•
	Adjudication/litigation	•		•	•
Hassle costs		•		•	•
Indirect costs					
Indirect compliance costs			•	•	
Other indirect costs	Offsetting/substitution effects		•	•	
	Transaction costs		•	•	
	Opportunity costs		•	•	
	Reduced competition		•	•	
	Reduced market access		•	•	
	Reduced investment/innovation		•	•	

***Citizens**: means citizens and society as a whole and refers to impacts that are widespread and do not affect a particular sub-group in a specific way;*
***Consumers** refers to a specific product or service. Consumers do not necessarily overlap with citizens but may be a sub-group e.g. a group of citizens targeted by a given regulation;*
***Business** includes all types of businesses including SMEs;*
***Public administrations** are EU, national, regional or local administrations.*

5. DIFFERENT TYPES OF BENEFITS

Benefits are often the objective of the initiative. They can be more difficult to classify as sometimes the beneficiaries can be a society-wide and more diffuse group than the stakeholders that bear the costs (see Table 1):

(1) ***Direct regulatory benefits*** – these are usually expressed as objectives of the intervention and can be broadly divided into two categories:

- The general *increase in welfare* of individuals and society by means of improvements in health, environment, and safety (also known as *non-market benefits*); and
- *Market efficiency improvements*, which include higher economic productivity, improved allocation of resources, removal of regulatory or market failures or cost savings but also information availability and enhanced product and service variety and quality for end consumers.

(2) ***Indirect regulatory benefits*** – similarly to indirect costs, indirect benefits occur as secondary impacts to the intended achievements of the intervention, i.e. they are not

explicitly formulated as objectives. They are often referred to as co-benefits (or ancillary benefits).They can be in general classified in to the following:

- *Spill-over effects* related to third-party compliance with legal rules (so-called '*indirect compliance benefits*'). These are benefits which accrue to individuals or business that are not the direct addressees of the initiative but who enjoy positive effects due to the compliance of others who are directly addressed (e.g. lower societal healthcare costs due to strategies to reduce obesity or tobacco smoking; suppliers of environmental technologies);
- *Wider macroeconomic benefits,* including higher GDP, productivity enhancements, greater employment rates, improved job quality etc.; and
- *Other non-monetary benefits,* such as protection of fundamental rights, social cohesion, reduced gender discrimination, international and national stability, etc.

The 'aggregate impacts' of a policy intervention: - All regulations usually aim, as an ultimate impact, to achieve some advancement in social welfare, which can be described in terms of efficiency or in other terms. These ultimate impacts encompass welfare, environmental quality, and more economic goals such as GDP growth and employment.

TOOL #57. METHODS TO ASSESS COSTS AND BENEFITS

Different methodological approaches can be used to estimate costs and benefits ex ante (impact assessment) or ex post (retrospective evaluation/fitness check). The most appropriate choice will depend on various factors such as the nature of the initiative, proportionality, and data availability. Box 1 considers a few important general issues:

Box 1. Important issues with respect to assessing costs and benefits

When assessing costs or benefits it is important to:

- distinguish between private or social, direct or indirect, and one-off or recurrent costs / benefits;
- avoid double-counting costs and benefits of regulation by recognising that the gains of one category and the losses of another may be two sides of the same coin[815];
- all costs (and benefits) generated by a new legal provision are by definition incremental costs, *i.e.* they are additional to the existing costs and benefits in absence of a new policy measure ('business as usual', BAU).

Costs

1. HOW TO ASSESS COMPLIANCE COSTS

Direct costs are those costs linked to the needs to divert resources to implement or abide by the requirements of a regulatory option[816]. **An important category of direct costs are the so-called compliance costs,** i.e. those costs incurred by businesses and other parties in undertaking the actions necessary to comply with the new regulatory requirements[817].

Compliance costs are often the aggregate of all direct costs generated by legislation: over time, they have become the subject of specific assessment methods in various countries. It is often useful to analyse (and estimate) compliance costs based on their individual components. Box 2 gives an overview of different compliance cost components.

It is important to distinguish the costs borne only once, usually upfront ('one-off' costs) from the costs that occur several times (recurrent costs) throughout the period in which the legislation/intervention remains in place. For the latter, it is necessary to identify the frequency with which the costs have to be borne (e.g. (bi)annualy, quarterly, monthly) and – in case of applying quantitative methods – to discount them to express them in the present value of money (see Tool #63 (*Cost-benefit analysis*) and Tool #64 (*Discount factors*)).

Direct costs linked to a regulatory initiative, will not only change the actions of the upstream and downstream businesses or other parties that are affected; but due to the economic

[815] For example, Economic impacts of EU clean air policies assessed in a CGE framework discuss an example for the European Clean Air Package. Table 4 illustrates that the direct abatement costs can be larger than the economy-wide costs (GDP), as the costs imposed to the polluting industries lead to increased production and employment in the sectors providing the green technologies.

[816] Assessing the Costs and Benefits of Regulation; CEPS (2013) a study prepared for the European Commission.

[817] OECD (2014), OECD Regulatory Compliance Cost Assessment Guidance. See page 62 for a list of regulatory compliance activities.

integration may lead to (sometimes significant) indirect costs or benefits to other parties. These indirect effects are, typically, measured by models such as general equilibrium models (see Tool #61 (*Simulation models*)).

Box 2. Compliance cost components	
Regulatory charges	
Regulation often affects businesses and consumers by imposing the payment of fees, levies, or taxes on certain stakeholders. These costs are often easy to calculate, as their extent is known. It is more difficult to assess who will bear those costs, as this might depend on the extent to which these costs are passed on to entities other than those targeted by the legal rule[818]. For example, copyright levies might be passed on downstream on end consumers in the form of higher prices for certain hardware devices.	
Administrative costs	
These are the costs of complying with administrative requirements imposed by regulation. For details, see Tool #58 (*EU Standard Cost Model*).	
Adjustment costs	
These are the incremental (i.e. non-business as usual) costs of complying with new regulation other than charges and administrative costs. They can be *broken down into*[819]*:*	
Implementation costs	The costs regulated entities incur in familiarising themselves with new or amended regulatory compliance obligations, developing compliance strategies, and allocating responsibilities for completing compliance-related tasks. In large part, therefore, they are short-term one-off costs.
Direct labour costs	The costs of staff time devoted to completing the activities required to achieve regulatory compliance. Only the costs of staff directly involved in undertaking these activities should be included: the costs of staff supervision/management are included in the overhead cost category (see below). Direct labour costs include two main elements: the cost of wages paid and non-wage labour costs.
Overheads	The costs of rent, office equipment, utilities and other inputs used by staff engaged in regulatory compliance activities, as well as corporate overheads, such as management inputs, that are attributable to compliance activities.
Equipment costs	Those costs incurred by businesses whenever they need to purchase items of capital equipment to comply with a regulation. This can include both machinery (e.g. equipment to treat the emissions from a production facility to conform to new emissions standards) and software (e.g. programs required to undertake real-time monitoring of actual emissions).
Material costs	The incremental costs incurred in changing (or substituting) some of the material inputs used in the production process to ensure regulatory compliance (thus, they are sometimes called 'input costs'). They are therefore recurrent costs.

[818] In economic theory, this ability to pass on part of the costs is defined by the elasticity of demand.
[819] The categorization proposed in the OECD Regulatory Compliance Cost Assessment Guidance.

Cost of external services	The cash cost of payments made to external suppliers providing assistance in achieving regulatory compliance. For example, faced with more stringent emissions controls, a firm may hire consulting engineers to advise on the available means of reaching compliance and their relative costs and benefits.

The extent to which different types of compliance costs are assessed can be guided by different principles. All have their limits and elements to consider are:

- The expected magnitude of compliance costs: the higher the expected cost, the more resources should be invested in estimating compliance costs;
- The distributional impact: the more heterogeneous the affected actors are or the more disproportionate the impacts are, the higher the need for a thorough the analysis;
- The availability of data: the greater the availability of data, the more comprehensive compliance costs should be quantitatively estimated;
- The nature of the initiative: when compliance can be broken down into a relative precise set of activities, total compliance costs can more easily be estimated by adding up the various costs of these activities for a typical business/citizen/public authority. Conversely, a top-down approach may be more appropriate in the case of complex policy proposals, where the range of starting positions across regulated entities is wide and/or where there are different ways to achieve compliance.

Making methodological choices, but also applying the principle of proportionate analysis, include seeking compromises by focussing on major cost drivers and relying on simplified assumptions (e.g. extrapolating data from some economic actors or member states to others). All key methodological choices and limitations need to be justified. When appropriate, a sensitivity analysis may be considered (See Tool #65 (*Uncertainty and sensitivity analysis*)). Box 3 gives an overview of various methods to estimate compliance cost components.

Box 3. Methods to estimate compliance cost components
Methods to assess compliance costs build up on the standard cost model[820], which is used specifically to estimate administrative costs. The following provides a short summary.
Regulatory charges
Charges = Targeted population x Frequency x Cost
For example, if 2 500 enterprises pay a licence fee of €500 twice a year, the total on a yearly basis will be (2 500 x 500 x 2) = €2.5 million.
Administrative costs
Administrative costs are the costs incurred by enterprises, the voluntary sector, public authorities, and citizens in meeting administrative obligations towards public authorities or private parties. Administrative obligations in a broad sense include labelling, reporting,

[820] See CEPS (2013) and OECD (2014) for references to existing methods which extend the standard cost model from administrative cost to compliance costs more generally.

registration, monitoring, and assessment needed to provide the information. In some cases, the information must be transferred to public authorities or private parties. In others, it only must be available for inspection or supply on request.

Whenever a measure is likely to impose significant administrative costs on business, citizens or public authorities, the EU Standard Cost Model presented in Tool #58 should be applied.

In the context of the reporting requirements linked to the 'one in, one out' approach, new and/or removed administrative costs for businesses should be encoded in an online OIOO calculator[821] established specifically for this purpose.

Adjustment costs

(1) Identify substantive requirements

These are all the activities necessary to comply with, except for those linked to administrative obligations and charges. Here one should distinguish between one-off and recurrent requirements.

(2) Estimate the population of stakeholders that must comply with each substantive requirement for each of the alternative options.

(3) Estimate the time needed for a 'normally efficient business', an 'ordinary citizen' or a 'normally efficient administration' to comply with each substantive requirement.

This stage is required to assess the cost of labour needed to comply with the substantive requirements. The concept of 'normal efficiency' is needed in order not to factor into the analysis the inefficiency of some of the targeted entities. This implies the assumption that regulated entities handle their substantive tasks neither better nor worse than may be reasonably expected. Information on the time it takes for normally efficient entity is usually gathered via surveys, interviews or drawing parallels with similar tasks for which the average time is already known.

(4) Estimate the 'BAU factor' for each substantive requirement and each of the alternatives, based on direct assessment or empirical data.

Estimation of the business-as-usual (BAU) factor is needed to estimate the incremental costs of regulation. It is often obtained by consulting targeted stakeholders or experts: BAU estimates involve the exercise of identifying costs that would normally exist regardless of new policy interventions. In some cases, the BAU factor can be estimated directly by looking at the share of costs associated with a substantive requirement that are borne by similar entities that are not targeted by specific legislative provisions.

(5) Consider segmenting the population by creating case groups differentiated according to size (micro, small, medium, large enterprises) or other dimensions (level of government for public administrations, availability of Internet connection for citizens, etc.).

This step will potentially identify any distributional impacts of costs imposed by the considered options and allow for designing mitigation measures if necessary.

(6) Estimate the adjustment cost associated with each substantive requirement for each segment and each alternative.

[821] The online OIOO calculator

The costs of each of the adjustment cost components (see Box 2 above) need to be estimated and summed up considering potential differences between segments. The calculation has to reflect that some of the costs will be one-off and some recurrent (e.g. cost of materials). For the latter, if they are monetised, discounting them is necessary to calculate their net present value (see Tool #64 (*Discount factors*))[822].

(7) Assess whether adjustment costs are likely to change over the life of the proposed legislation.

One should assess whether, because of entry/exit of businesses, technological innovation, 'learning by doing' or any other relevant factor, the impact of the costs identified is likely to change over time. This should be considered in a prospective analysis on regulatory costs, and – if possible – coupled with sensitivity analysis on the assumptions behind the evolution of costs over time.

(8) Calculate all incremental adjustment costs and extrapolate them to reach a total estimate for each of the alternative options considered.

The incremental adjustment cost is equal to the calculated adjustment cost minus the business-as-usual cost. The incremental adjustment costs should then be extrapolated to the whole population of affected entities per identified segment.

The accuracy of these methods depends significantly on the extent to which resources are devoted to data collection and prospective analysis. **Without reliable data and/or robust assumptions on future evolutions, results of an assessment of costs and benefits can only be considered as indicative of the relative magnitude of compliance costs across different alternative options.**

2. HOW TO ASSESS ENFORCEMENT COSTS

Enforcement costs are those costs direct borne by public authorities in implementing, administering, and enforcing regulatory requirements.

They can include the cost of publicising new requirements, establishing licencing or permit systems, dealing with queries and applications, handling complaints, implementing inspections and audits to verify compliance and sanctioning non-compliance[823].

In principle, enforcement costs can be estimated following a similar bottom-up approach to the one described in for adjustment costs: first defining the activities required to implement and enforce legislation, then estimating their frequency and their cost taking into consideration the BAU factor and possibly distinguishing between different case groups as appropriate. Enforcement authorities may be in a position to provide good unit cost estimates for different types of activities.

In reality, however, estimating these costs ex ante at the stage of Commission proposal may be complex. First, data are rarely available. Second, implementation and enforcement activities often cannot be defined (and thus costed) since they are to be decided and implemented by Member States at a later stage. However, Member States may be in a position to provide estimates of costs for similar activities performed by them in the context of a similar enforcement area. In this case, one could complement the estimate with a

[822] Further useful guidance on this can be found in chapter 3 of the OECD (2014).
[823] For a list of possible implementation and enforcement activities, see p. 63 in OECD (2013).

qualitative assessment in order to consider any trade-offs and synergies between business (or citizens) adjustment costs and enforcement costs[824].

3. HOW TO ASSESS CUMULATIVE COSTS

Every new policy proposal should be assessed on its own merits. For this reason, impacts are assessed against a baseline, meaning that only the incremental costs and benefits of the new initiative need to be estimated. It is the sign and the magnitude of the net change in costs and benefits that matter most for the proposal of new policy initiatives.

Cumulative cost assessments (CCA), however, and look at the aggregate (or cumulative) costs of all different regulations from the point of view of a single economic sector. The exercise is of backward-looking nature and can be used as an input to define the baseline scenario. **It is a partial approach, which by definition does not look at benefits.**

A prominent tool for making CCAs is the Dutch Compliance Cost Assessment tool ('CAR model') which was conceived essentially for retrospective analyses of existing legislation[825].

Not all regulatory proposals lead to direct cost increases. At times, the very aim of a regulatory proposal is to reduce existing, cumulative regulatory costs either by simplifying existing EU legislation or by harmonising regulations across Member States and thus generally reducing compliance costs for businesses operating across the single market. In case of reducing administrative costs, they should be recorded in a dedicated online calculator supporting the application of the 'one in, one out' approach (see Tool #58 (*EU Standard Cost Model*)). **It is recommended that all revisions of legislative acts (i.e. REFIT initiatives) have simplification as one of its specific objectives where this is relevant**. Please see Tool #2 (*The regulatory fitness programme (REFIT) and the Fit for Future Platform*) and Tool #15 (*How to set objectives*) for further details.

It is advisable to consider cumulative impacts to the extent that this may be possible and proportionate. Some reflections in this regard can be found in Box 4.

Box 4. Assessing cumulative impacts	
Why?	Because it helps avoiding redundant requirements (for instance, reporting ones) and/or highlights opportunities to simplify legislation. Cumulative assessments can also help in defining better the baseline scenario.
	Because a good assessment of indirect impacts may depend upon a good understanding of cumulative impacts. For instance, the impacts on sectoral competitiveness of an increase in regulatory costs depend upon the overall aggregated cost structure (incl. cumulative

[824] Thus, an option that provides greater flexibility in the ways in which business can comply with the regulatory requirements may minimise costs to firms, but may increase the costs of administering the regulation, since verifying compliance will be more complex and involve a higher degree of professional judgement. Total direct costs may well be higher than under a less flexible regulatory option. Total costs would of course also depend on indirect impacts such as impacts on business competitiveness, innovation, the ultimate goal of the regulation, etc.

[825] The Study on Assessing the costs and benefits of Regulation prepared by the Centre for European Policy Studies (December 2013); pp 70 for description of the model and its strengths and weaknesses in relation to impact assessments and evaluations.

	regulatory costs) of the sector vis-à-vis international competitors.
How and when?	During public consultations when stakeholders discuss interactions between a proposed initiative and the existing body of legislation.
	When conducting fitness checks as part of the input to assess efficiency of (all or part of) the legislation affecting a specific sector.
	When designing policy options, the lead service and the ISG should cross check the proposal with the existing body of legislation (across the sector and policy areas) for possible redundant requirements, overlaps etc. In doing so, the results of the ex-post evaluation(s) may provide useful information.
	When assessing impacts and, notably indirect impacts and impacts on micro and small enterprises[826]. No generally recognised standard methodology exists for the consideration of cumulative impacts. However, a growing number of studies are generating data by sector and firm type. The methodologies can also inform quantitative models.

Benefits

4. HOW TO ASSESS COST SAVINGS

Box 5 gives an overview of different methods to estimate cost savings.

Box 5. Methods to estimate direct cost savings
Whenever a policy option leads to a reduction in regulatory charges, one could follow the same approach as suggested in Box 3 to estimate the value of the reduction.
Whenever a policy option leads to a reduction in adjustment and administrative costs, the same approach is applicable as in Box 3 of this tool and in Tool #58 (*EU Standard Cost Model*) on administrative costs to estimate the value of the reduction.
Whenever a policy option leads to a reduction in enforcement costs, to estimate the value of the reduction (or at least assess its relevance qualitatively), instructions given in section 2 of this tool are valid.

All the usual caveats applying to bottom-up estimation approaches apply to the above. In particular, it is very important to complement any such estimation with an assessment of indirect costs and of direct and indirect benefits. This guarantees that cost savings do not reduce regulatory benefits (or at least do not reduce them in a manner that worsens both the effectiveness and the efficiency of an existing policy). It is also necessary to consider possible trade-offs among different categories of costs. Box 6 lists such trade-offs using administrative obligations as an example.

[826] See Tool #23 (*The 'SME test'*)

Box 6. Verifying the effective nature of cost savings

A proposal may reduce administrative costs, but at the same time increase other compliance costs for the same targeted businesses.

Administrative costs (ACs) constitute only a subset of costs imposed on businesses by legislative acts. For example, the implementation of an e-government or any other IT-enabled solution can reduce the amount of time related to compliance with the information obligation. At the same time, however, it may require a degree of investment in upgraded IT equipment and training of employees, which would not be considered as ACs, but fall generally in the category of adjustment costs. Similarly, a proposal that reduces ACs may increase public expenditure in monitoring and enforcement (see below): these costs may be recovered by the government through higher tax burdens, thus increasing direct charges. Finally, a proposal may reduce administrative obligations by requiring structural changes in the production process, which would guarantee a certain level of product safety without any need for burdensome certifications: in this case too, ACs are reduced, but adjustment costs may increase.

A proposal may reduce administrative costs, but at the same time increase administrative costs of a different origin.

With multi-level governance, the reduction of ACs achieved by eliminating some information obligations at e.g. the EU level may require the introduction of new information obligations at e.g. the national or regional levels.

A proposal may reduce administrative costs, but at the same time increase costs for other private actors (businesses and/or citizens, workers).

For example, reducing labelling obligations for products may increase information costs borne by consumers, who would need to collect their information from other sources in order to make an informed choice of what products are most likely to fit their preferences.

A proposal may reduce administrative costs, but at the same time increase monitoring and enforcement costs for public authorities.

This is often the case whenever the eliminated information obligations involve the keeping and reporting of information available to businesses, but not to public authorities. For example, the reporting on respecting the hygiene standards or on the large exposures by banks are typical instances of highly burdensome activities for businesses. These information obligations are vital for public authorities. Without them, public authorities would have to deploy more resources to obtain the information, which is likely to lead to more inspections and enforcement costs – in the two examples, more hygiene inspections and more investigations into the riskiness of banks' exposure *vis à vis* certain clients.

A specific case of savings can occur whenever there are options that have an impact on the **Single Market**, especially when such options entail the harmonisation of national legislation. Box 7 presents the specificities of such a category of cost savings.

Box 7. Single market cost savings

Savings might emerge whenever EU legislation harmonises fragmented and inconsistent national legislation. When legislation is fragmented across Member States, companies wishing to engage in cross-border trade have to incur 'adaptation costs', such as:

- ***Having to change contracts or other practices to comply with differing national legislation.*** Monetising these costs is normally possible. One way of doing it is to collect data directly from companies and validate them with experts.
- ***Having to modify standards or equipment, or train personnel to deal with differing national legislative requirements.*** These costs are easily monetised by referring to market prices, and (in the case of equipment) depreciating these assets over time.
- ***Incur additional administrative costs because national legislation contains different information obligations***, which have to be complied with and which would not be incurred if the company refrained from entering the national market. In this case, one has to estimate the time that would be spent complying with the additional information obligations and convert this into a monetary value by using data on labour costs for the job profile of the person that would perform the relevant administrative activities.

A number of *caveats* should be kept in mind when performing these calculations. First, adaptation costs might not be incurred by companies if they keep internal compliance programs that apply to one or more countries: for example, if a company adopts an internal antitrust compliance program that is tailored to the most restrictive country, this will automatically mean that the company also complies with legislation in less strict countries. Also, the magnitude of administrative burdens should be gauged against the 'BAU factor', *i.e.* the extent to which the activities performed to comply with national legislation would be performed anyway, even when not required by EU law.

Finally, cost savings are only one category of benefits one has to deal with when looking at harmonisation of legislation. Indirect benefits may emerge due to market efficiency impacts. Estimating these benefits is normally not prohibitively difficult but for accurate monetisation, one needs data on demand and supply elasticities[827]. Expert guidance for this type of estimation[828] may be helpful (See Tool #25 on impacts in the internal market).

5. NON-MARKET BENEFITS

Assessing quantifiable direct or indirect benefits does not substantially differ from assessing the corresponding costs and may use the similar methods. Wider economic benefits, like GDP growth, employment increase, or productivity growth, can be assessed, among others, with modelling (Tool #61 (*Simulation models*)). Social cost/benefit analysis assesses the net value of a policy or project to society. Many non-market benefits (e.g. health, quality of the environment) are often expressed in physical units. Monetisation of non-market benefits is easier when the values can be linked to market prices. E.g. air pollution might reduce crop yields, thus allowing for relatively straightforward monetisation. Other non-monetary

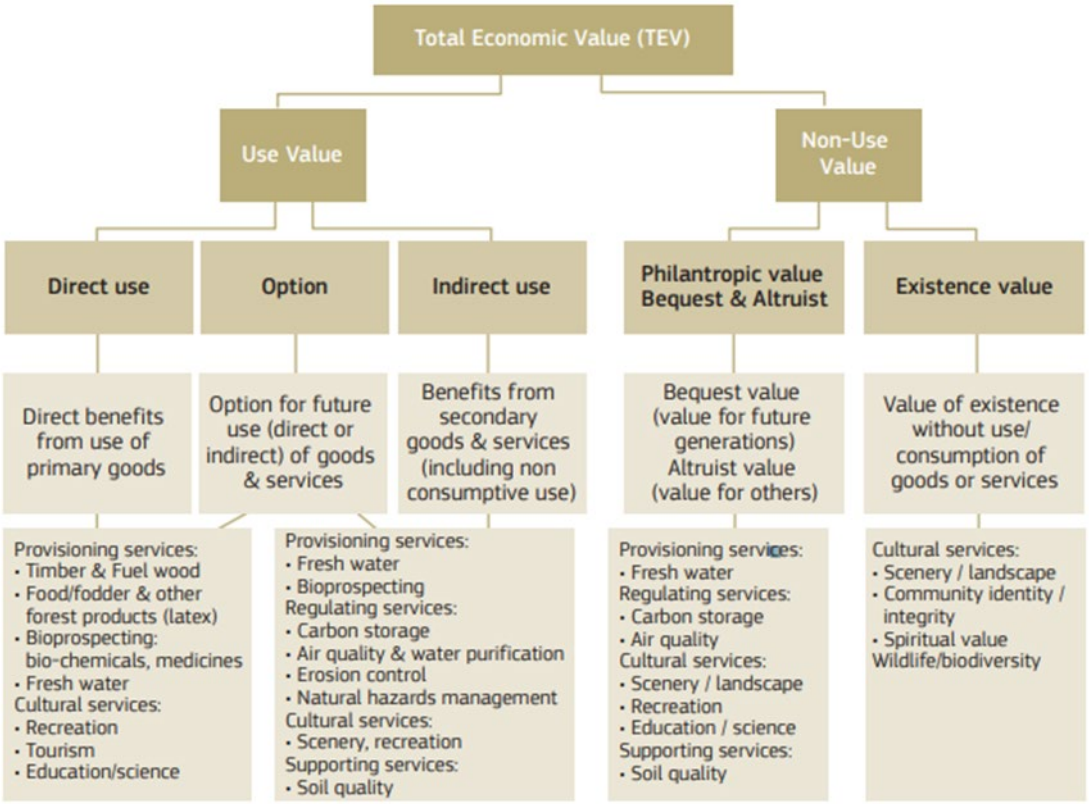
[827] In the US, dedicated databases make it easier to estimate the response of supply and demand to a change in price or in the quality of products (e.g. here, in particular the section on commodities and food elasticities). In the economic literature, several estimates of elasticity are available, which could be collected into a single dataset. See for example in relation to air transport.

[828] For a general presentation see CEPS (2013), p. 178-182.

benefits, such as improvements in protection of fundamental rights, social cohesion, or international stability, are less straightforward to measure and are assessed by surveys or proxy indicators (e.g. counting LGBTQ laws). However, the full value of many goods (benefits) such as health, environment, or education cannot be easily deducted from the market price. However, these important social impacts cannot be ignored in policymaking. The valuation of non-market impacts is challenging but could be undertaken wherever possible. Complementary tools exist to compare the merits of policy options where monetary information may be limited (e.g. multi-criteria decision analysis, Tool #62). Box 8 looks at dedicated methods attempt to assign market values and monetary units to these benefits.

Box 8. Market based approaches: Stated preference and Revealed preference

Valuation techniques allow for estimation of the total economic value (TEV), which refers to the value derived by people from a natural resource even if the services that it provides are not fully reflected in market prices. It is an aggregation of the (main function based) values provided by a given ecosystem[829].



Market based approaches estimate the change in utility by simulating a market. The willingness to pay (WTP) is the maximum amount of money an individual is willing to give up in order to receive a good. The willingness to accept (WTA) is the minimum amount of money they would need to forego or give up a good. The willingness to pay/accept depends on their income level so valuations are usually obtained by averaging across income groups. The approaches can be used to estimate the value of improved health outcomes[830, 831] (such as

[829] Economic Benefits of Natura 2000 report
[830] See Tool #32 (*Health impacts*)

from reduced air pollution), reduced road congestion, reduced road fatalities and injuries, disamenity (e.g. from waste disposal and quarrying) and recreational amenity (e.g. forests).

'Revealed preference' vs 'stated preferences' vs 'experiments'[832]

- Revealed preference techniques involve inferring the implicit price placed on a good by consumers by examining their behaviour in a similar or related market. For example, the value of house prices and its relationship to ambient noise or the travel costs incurred by individuals who wish to enjoy a forest or other recreational site.

- Stated preference techniques use specially constructed questionnaires to describe a hypothetical choice within a hypothetical market in order to elicit estimates of the willingness to pay or willingness to accept. When using stated preferences, the main choice is between contingent valuation and choice modelling. The former elicits WTP or WTA via direct questions on the amounts they would be prepared to pay to receive a particular good while the latter present respondents with a series of alternatives and asking for their preference.

- Experiments are different to revealed or stated preference surveys, as subjects in experiments make incentivised choices, and may accrue benefits and incur losses. In revealed and stated preference surveys, these types of incentives are not present. (see Tool #69 (*Emerging methods and policy instruments*))

While revealed preference methods are perceived to be more reliable, these techniques cannot be used for new assets or new users. Here, stated preference techniques can be an alternative. It may be difficult to judge the reliability of estimates emerging from a single study using a single method. In any event, a range of values could indicate the sensitivity.

In the absence of an existing reliable and accurate monetary valuation of an impact, a decision could be made whether to commission a study, and if so, how much resource to allocate to the project in line with the principle of proportionate analysis. Factors to consider include (i) whether further research is likely to yield a robust valuation; (ii) whether the results will be useful for future IAs; (iii) how accurate the valuations need to be; (iv) the political importance/magnitude of the policy initiative and the expected impacts.

The technique of 'benefits transfer' (or 'cost transfer') can also be used to estimate values of impacts that do not have market prices. To safe times and resources, values obtained in one study are transferred to a different study. For example, estimates of the costs of preventing a motorway accident in one Member State might be used to estimate the costs in other Member States. Using this technique increases the uncertainty of the estimated values but can be helpful to give an order of magnitude of likely impacts, even with time or money constraints. A good example of benefits transfer is the DG MOVE Handbook on the **external costs of transport** (Version 2019)[833] that provides information on how to generate state-of-the-art

[831] For example, the OECD has done a study to explain the variation in the estimates of Value of Statistical Life (VSL) estimates according to the characteristics of risk (type and size of risk, baseline risks, latency, etc.), and socio-economic characteristics (age, income, gender, health status, etc.). Their meta-analysis suggests a VSL of around 3.5 million euro. In the chemicals context, standard values are also found for different health end points.

[832] See overview in Table 29 of the CEPS 2013 study on Assessing the Costs and Benefits of Regulation, pp. 185. http://ec.europa.eu/smartregulation/impact/commission_guidelines/docs/131210_cba_study_sg_final.pdf

[833] https://ec.europa.eu/transport/sites/transport/files/studies/internalisation-handbook-isbn-978-92-79-96917-1.pdf

estimates for all main external costs of transport. The European Environment Agency has made accessible some unit values for the valuation of **industrial activity**. It is also difficult to carry out benefits transfer for biodiversity and water pollution or scarcity. Some environmental issues are highly location specific (e.g. biodiversity, water pollution), and, hence, benefit transfer may be difficult. The environmental valuation reference inventory (EVRI) database[834] may help with case studies.

When valuing impacts, the proportionality principle applies, as in all parts of impact assessment: it may not be worth applying these techniques for very small impacts. Further, there may be significant impacts that cannot sensibly be monetised, and these should be presented in non-monetary units (e.g. air pollutant concentrations) or in more qualitative terms[835]. Where costs cannot be valued in monetary terms, they should still be reflected in the impact assessment.

6. OTHER METHODS USED FOR ESTIMATING COSTS AND BENEFITS

Simulation models (economic or physical), such as partial or general equilibrium models, are commonly used in the preparation of impact assessments. The JRC has established a Competence centre on modelling (CC-MOD) which can provide information about the available models and help and assistance in using such models (see Tool #61 (*Simulation models*)).

The tools on cost-benefit analysis and multi-criteria decision analysis discuss how the costs compare to benefits. Box 9 gives an overview of other analytical methods that do this comparison.

Box 9. Other analytical methods

A **SWOT analysis** is used to identify the Strengths, Weaknesses, Opportunities and Threats in relation to a project/organisation and how such an assessment will change over time.

In the context of an evaluation, this method can be used for assessing the services provided by a project/programme. SWOT can take past weaknesses and transform them into a constructive learning process. SWOT is not an analytical tool by itself; instead it is a way to synthesise preceding analyses and use them for developing a strategy.

Least cost analysis is primarily used in the impact assessment context. It only looks at costs, in order to select the alternative option that entails the lowest net cost. You should choose this method whenever benefits are fixed, and you only need to choose how to achieve them. Costs do not need to be precisely monetised or even quantified but their relative magnitude across options should be determined.

Cost-effectiveness analysis (CEA) entails that you quantify (not monetise) the benefits that would be generated by one euro of costs imposed on society. While CEA is closely related to CBA, instead of monetised benefits it uses other measures, such as increased life expectancy, educational attainment, abated emissions. A typical method to compare options for an impacts assessment is the so-called **benefit-cost ratio**, which means dividing the benefits by

[834] https://evri.ca/en/content/about-evri
[835] https://www.oecd.org/env/tools-evaluation/CBA-brochure-web.pdf

costs. This method can be used for expenditure programs, as it helps to identifying the 'value for money'. In the evaluation context, the cost-effectiveness analysis can compare the evaluated intervention against best practice or similar interventions. It can also be used to assess effectiveness of the implementation process where different implementation approaches have been pursued. CEA is less easily applicable to interventions with more than one main objective. If the intervention aims to achieve a number of objectives (e.g. job creation and environmental protection), or have indirect impacts, the results of CEA may be misleading or irrelevant.

A **counterfactual analysis** is a statistical method devoted to quantifying the size of the effect of a given intervention (see Tool #46 (*Designing the evaluation*))

7. FURTHER INFORMATION

- Study prepared by the Centre of European Policy Studies on the assessing the costs and benefits of regulation (2013).
- OECD Regulatory Compliance Cost Assessment Guidance (2014).

TOOL #58. EU STANDARD COST MODEL

1. ADMINISTRATIVE COSTS AND THE 'STANDARD COST MODEL'

Administrative costs are a specific type of compliance costs incurred by enterprises, public authorities, and citizens **in meeting administrative obligations**. This captures a broad range of administrative activities including labelling, reporting, registration, provision of data, as well as monitoring and assessments needed to generate the information. In some cases, the information has to be transferred to public authorities or private parties. In others, it only has to be available for inspection or supplied on request.

Box 1. Example of administrative costs vs. adjustment costs

- A regulation on air quality sets an obligation to keep a register of pollutant emissions and an obligation to meet an air quality threshold.
- Keeping a register of pollutant emissions is an administrative cost, while action taken to meet an air quality threshold is not. That type of compliance cost is referred to as 'adjustment cost' because the obligation affects the core industrial or economic activity.
- Keeping a register does not entail in itself any obligation to change the production process, the nature of the end-products or the treatment of emissions. Meeting the pollution threshold will require a substantive change to the industrial activity (for instance the installation of new filters).

Whenever a measure is likely to impose significant administrative costs on business, citizens or public authorities, the **EU Standard Cost Model** should be applied to the extent that the underlying data is available. The main aim of the model is to assess the **net cost of administrative obligations** imposed by EU legislation[836]. The results of this analysis should be presented in **Annex 3** ('who is affected and how') of the impact assessment or for evaluations or fitness checks in **Annex IV** ('overview of benefits and costs'). In case an initiative imposes new or eliminates existing burdens on businesses or citizens, these should be measured with a dedicated online calculator[837], supporting the implementation of the 'one in, one out' approach (see Tool #59 (*Cost estimates and the 'one in, one out' approach*)).

2. OUTLINE OF THE MODEL

2.1. Definition of administrative costs and administrative burden

Both **recurring administrative costs** and, where significant, **one-off administrative costs** have to be considered.

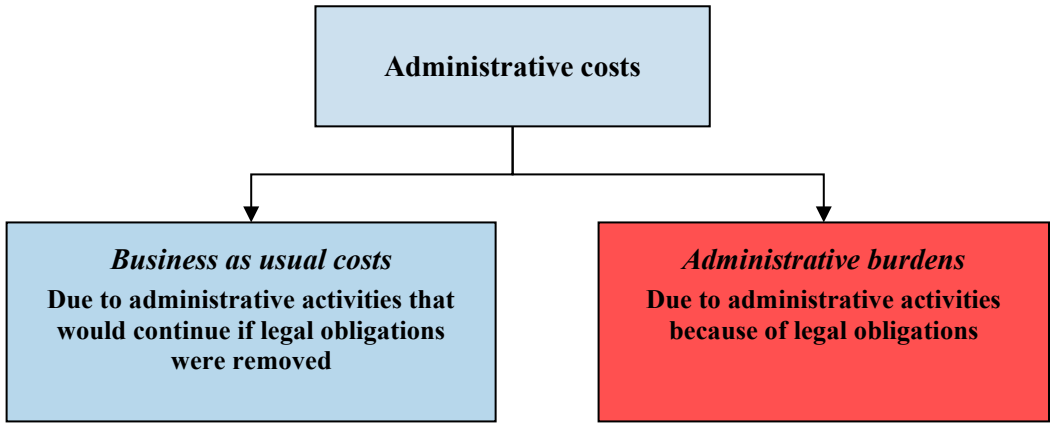
The administrative costs consist of two different cost components: the **business-as-usual costs** and **administrative burdens**. While the business-as-usual costs correspond to the costs resulting from collecting and processing information that would even be done in the absence

[836] Net costs = costs introduced by a proposal if adopted, minus the costs it eliminates at EU and/or (sub)national level.

[837] The online OIOO calculator

of any legislation, **administrative burdens stem from the part of the process which is done solely because of a legal obligation**[838].

Box 2. Components of administrative costs



2.2. Core equation of the cost model

Administrative net costs are assessed by multiplying the **average cost of the required administrative activity** (Price) with the **total number of activities performed per year** (Quantity) and by subtracting the cost of administrative activities removed (at EU/national level).

The average cost per activity can be estimated by multiplying a tariff (based on average labour cost per hour including pro-rated overheads) and the time required per action. Where appropriate, other types of costs such as outsourcing, equipment or supplies' costs should be considered[839]. The quantity will be calculated as the frequency of required actions multiplied by the number of entities concerned. In case of multiple relevant administrative activities per information obligation these need to be summed up to calculate the administrative cost per information obligation. The core equation of the EU Standard Cost Model is as follows:

Box 3. Core equation of the EU Standard Cost Model

$$\text{Administrative cost} = \Sigma\, P_N \times Q_N - \Sigma\, P_R \times Q_R$$

where **P** (for Price) = Tariff x Time; **Q** (for Quantity) = Number of businesses x Frequency; and N – new obligations, R – removed obligations at EU/national level

[838] Most businesses would for instance have an accounting system, even in the absence of legal bookkeeping, but would not necessarily provide caloric value information for all their products.

[839] Many small businesses, for instance, use services by external accountants, chambers of commerce, and professional associations to fulfil certain information obligations set by Company Law.

2.3. Scope of application of the model and expected level of accuracy

The effort of assessment should remain proportionate to the scale of the administrative costs imposed by the legislation and should be determined according to the principle of proportionate analysis (see Tool #12 (*How to apply proportionality to impact assessments*)). Therefore, there is no need to assess the administrative costs when these are **bound to be insignificant,** for instance, when **little equipment** is required**, if the amount of time per action is small** and **the frequency low**. Such decisions (i.e. no costing) should be taken on a case-by-case basis and should be justified. In order to keep assessment of costs at a reasonable level and ensure comparability of results, **estimates will be based on standard assumptions** simplifying the complex reality of the EU. These assumptions are presented together with **step-specific guidelines below**.

3. STEP-BY-STEP GUIDE

The assessment of positive or negative effects on administrative burden on businesses, citizens or public administrations resulting from EU legislation should begin with a **full mapping of introducing new or eliminating existing administrative obligations** for each of the options under review. Such a comparative table will usually indicate the type of administrative obligation, its specific requirements, the target group, and the frequency of that obligation.

In the first phase, the mapping should clearly show how policy options differ in terms of administrative obligations, and, in particular, identify those obligations that are likely to impose **significant administrative burdens**. The significance (high – medium – low) is determined by a qualitative assessment of the likely number of entities concerned as well as the frequency and complexity of the required data.

Significant burdens will then be quantified (monetary estimates) on the basis of the **EU Standard Cost Model** (see Box 3).

A greater level of detail is expected for the monetary assessment of administrative burdens stemming from the preferred option. This detailed application can be divided in a number of steps. The entire workflow is summarised in Table 1, followed by a description of each step.

Table 1: Step by step application of the model

Phase I: Preparatory analysis	
Step 1	**Identification and classification of additional administrative obligations and obligations eliminated at EU/national level** (e.g. labelling of products) and its specific requirements (e.g. the label must provide the date of production and composition of the product)
Step 2	**Identification of required complementary actions** (e.g. training members and employees about the information obligations, filling forms)
Step 3	**Identification of target group(s), also called segmentation** (e.g. large enterprises that have to fulfil obligation 'A' and small enterprises that have to fulfil obligation 'B', the size of the enterprise being defined by its

	turnover)
Step 4	**Identification of the frequency of required actions** (e.g. enterprises have to fill a form once a year)
Step 5	**Identification of relevant cost parameters** (e.g. particular relevance of external costs – using accounting firms – and equipment) **Qualitative assessment of significant burdens** (i.e. applying a threshold test to determine which information obligations need to be quantified)
Step 6	**Choice of data sources** and, if necessary, development of data capture tool(s) (e.g. deciding that the number of entities concerned will be extrapolated on the basis of data available on Eurostat, but that the number of hours each need to perform required actions will be based on the results of interviews of enterprises; for the latter task, preparation of an interview guide and selection of a representative sample of entities including organisations representing citizens)
Phase II: Data capture and standardisation	
Step 7	**Assessment of the number of entities concerned** (e.g. 100 000 SMEs) (For businesses, this step requires the knowledge of the market structure in the regulated sector.)
Step 8	**Assessment of the performance of a 'normally efficient entity' in each target group** (e.g. enterprises spend once a year, on average, 25 hours of work by an engineer to gather information and 5 hours of work by a clerk to fill the annual form)
Phase III: Calculation and reporting	
Step 10	**Assessment of the 'business as usual' costs, extrapolation of data to EU level** (assessment of the proportion of the costs that would have been borne even if there was not obligation at the EU level, e.g. a company would have to present financial reports even if there were no EU rules on that and the cost of such reports could be similar)
Step 11	**a) Reporting in an impact assessment and – for business costs – in the 'one in, one out' online calculator.** **b) If no impact assessment is conducted due to political urgencies but should have been, then reporting is done in post-adoption document published within three months after the adoption.**

Step 1: Identification and classification of administrative obligations

To facilitate the assessment of administrative costs and to improve data analysis (identification and comparison of the most burdensome types of obligation across various

sectors), services are asked to use the following typology of the administrative obligations (Box 4) when inserting relevant information obligations in the dedicated online OIOO calculator for the reporting of administrative costs.

Box 4. Types of administrative obligations

- Notification of (specific) activities or events (e.g. for transportation of dangerous cargos; when an accident affects the environment)
- Submission of (recurring) reports (e.g. annual accounts)
- Information labelling for third parties (e.g. energy labelling of domestic appliances; price labelling)
- Non labelling information for third parties (e.g. financial prospectus)
- Application for individual authorisation or exemption i.e. obligation to fulfil each time a particular task has to be carried out (e.g. building permits; road transporters applying to be exempted from Sunday driving ban)
- Application for general authorisation or exemption (e.g. license granting permission to engage in an activity such as banking or liquor selling)
- Registration (e.g. entry in a business register or a professional list)
- Certification of products or processes, i.e. obligation to deliver a certificate (e.g. treatment facilities having to issue a certificate of destruction of a vehicle) or to get a certificate (e.g. aeronautical products and organisations involved in their design, production and maintenance must get the certification of the European Aviation Safety Agency – EASA)
  - *Inspection on behalf of public authorities (e.g. businesses having to monitor working and other conditions for employees, such as those related to occupational health and safety)**
  - *Cooperation with audits & inspection by public authorities or their appointees (e.g. obligation to cooperate with workplace inspection), including maintenance of appropriate records (e.g. obligation for treatment facilities to keep records about waste electronic equipment entering and leaving the treatment facility; obligation for hotels to keep a visitor log; these records must be presented during the inspection)**
  - *Application for subsidy or grant (e.g. to structural or cohesion funds)**
- Other

* This type of administrative cost is not subject to offsetting in the context of the 'one in, one out' approach (see also Tool #59 (*Cost estimates and the 'one in, one out' approach*))

Distinguishing an administrative obligation from other regulatory obligations is normally straightforward. There could however be a number of **borderline cases** where it is difficult to decide whether an obligation falls within the scope of the model or not. Ideally, such borderline cases are assessed to ensure consistency with respect to the decisions taken in other similar areas.

Box 5. Examples of borderline administrative obligations

- Costs induced by exercising a right to complain. These costs are not considered as an administrative cost by Member States quantifying administrative costs using the Standard Cost Model because there is no 'obligation' to complain.
- Costs induced by inspection. The usual purpose of an inspection is to collect the information needed to verify compliance with legal obligations (review of corporate books, etc.). Ensuing costs are clearly administrative costs. However, inspections are sometimes used to collect information unrelated to legal obligations (level of satisfaction of businesses, etc.). Submitting to such inspection is by definition voluntary and ensuing costs therefore fall outside the definition of administrative costs imposed by legislation.
- Costs induced by policy assessment. Some EU programmes require Member States to draw up national reform programmes. Designing a reform programme is quite different from an obligation to provide information. However, designing monitoring schemes, collecting data on the implementation of the policy, filling tables and submitting them to the Commission are clearly linked to information obligations. So policy design should not be considered as administrative cost, but the design of policy assessment should.
- Costs induced by the obligation of drawing safety plans. Some EU acts require businesses to design staged evacuation strategies, conduct exercises to verify that everyone knows what to do and when, etc. (e.g. plans for so-called Seveso establishments, airports, etc.). This is of course quite different from an obligation to provide information; resulting costs should therefore not be considered as administrative burden. The only eligible costs here basically are those linked to the obligation to collect information about impending risks (safety plans must often be based on a risk assessment) and the obligation to file and/or send the safety plan.
- Testing costs. When business have to submit their products & processes to a test in order to get an authorisation or a certificate, these testing costs are not considered as administrative costs.
- Some EU legislative acts and proposals also mention the possibility for Member States to ask for additional information (i.e. '*…Member States may … require the inclusion of other statements in the annual accounts in addition to the documents referred to in the first subparagraph …*'). Such possibilities are not to be included as EU administrative obligations, insofar as Member States are not obliged to ask that information. Nevertheless, such possibilities should be documented as they often pave the way for Member States' additions ('gold-plating').

Step 2. Identification of required complementary action

The services are asked to use the following typology on the type of required action.

Box 6. Types of required complementary action

- Familiarising with the information obligation
- Training members and employees about the information obligations
- Retrieving relevant information from existing data
- Adjusting existing data and producing new data
- Designing information material (e.g. leaflet conception)
- Filling forms and tables (including recordkeeping)
- Holding meetings (internal/external with an auditor, lawyer etc.)
- Inspecting and checking (including assistance to inspection by public authorities)
- Copying (reproducing reports, producing labels or leaflets)
- Submitting the information to the relevant authority (e.g. sending it to the relevant authority)
- Filing the information
- Buying (IT) equipment & supplies (e.g. labelling machines) to specifically used to fulfil information obligations
- Other

Step 3. Classification by regulatory origin

In order to enhance transparency on who is responsible for what, the **regulatory origin of administrative obligations** needs to be identified. In the context of the impact assessment, services are only requested to determine costs originating from the EU level, not those that may originate at national or lower levels. When the Commission assesses a possible measure, there is no point guessing what level of 'gold-plating' transposing authorities in each Member State might introduce. The Commission only has to account for proposals transposing international obligations in the EU and those resulting from its own initiative.

Two simple rules should be used for this:

- If the obligation arises entirely from an EU Regulation, i.e. needs implementation into the national legislative frameworks without any changes, 100% of costs induced by the obligation should be assessed.
- If the obligation is set in EU Directive, i.e. requires transposition by Member States and they do not limit themselves to what is needed to meet the obligation (by 'gold-plating'), attribute only the percentage of costs resulting from the obligation set in the EU legislation. 'Gold-plating' in the case of administrative obligations refers, among other things, to increasing the reporting frequency, to add 'data requirements' or to

widen the target groups. Naturally, only evaluations, not impact assessments, can assess the size of 'gold-plating'.

Step 4. Identification of target groups

As for the target groups, it may be useful to distinguish between groups on the basis of their size, type or location. Size may be particularly pertinent for enterprises. It is indeed often the case that an obligation is more burdensome for small enterprises than for large ones benefitting from economies of scale. Regulation often adjusts the type of administrative obligations according to a number of objective criteria (number of employees, turnover level, financial capacity of the citizens, etc.). A more in-depth knowledge of the market structure in the regulated sector is required to be able to properly identify the affected groups if they are not specifically targeted by the initiative.

Step 5. Identification of the frequency of required actions

The frequency indicates how many times per year an action is required[840].

In some cases, the frequency may vary in time. For instance, in a number of statistics regulations such as Intrastat[841], enterprises have to report if their dispatches are above a set threshold. Their level of intra-EU sales will therefore determine if they have to report or not. Here again, the advice is to keep things simple and proportional. If such fluctuations concern a limited number of enterprises, they should not be considered.

Step 6. Identification of relevant cost parameters

The relevant cost parameters are determined by the core equation of the standard cost model (see Box 3). It is assumed that the main costs induced by administrative obligations are **labour costs**. Where appropriate, **equipment or costs of supplies** per action should be considered or used as the basis for analysis (rather than taking time as the basis unit).

The cost parameters for the *price per action* (administrative action carried by the targeted entity itself) are the (i) **number of minutes spent on a specific action**, (ii) the **hourly pay** of those performing the action. This hourly pay should correspond to the gross salary plus overheads costs (25% by default)[842].

The cost parameters for *equipment & supplies* (i.e. acquired by the targeted entity to comply with the information obligation and solely used for that purpose) are the **acquisition price** and the **depreciation period** (service life of 'x' years)[843].

The cost parameters for the *outsourcing costs* (administrative action contracted out) are what the service provider charges on average per administrative obligation, per entity and per year.

---

[840] By definition, that notion does not apply to one-off costs such as 'familiarising with the information obligation'.

[841] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:l11011a

[842] The current tariffs are available in the dedicated online OIOO calculator for the reporting of administrative costs.

[843] For instance, barcode printer and scanner.

Step 7. Choice of data sources and, if necessary, development of data capture tool(s)

Data collection methods can differ for individual cases and include focus groups, consultation of stakeholders, field trials, consultancy studies, and expert assessment. Irrespective of the source and mode of collection, the collected data needs verification and interpretation.

In standard cases, it will be sufficient to produce overall estimates based on the method above as well as on available EU statistics;

In exceptional cases, **field work limited** to a sample of Member States and/or questionnaires sent to a standard sample of the business community or organisations representing individuals (for example, consumers), and simulation may have to be used.

Step 8. Assessment of the number of entities concerned

In order to ensure comparability of estimates made by different DGs, the assessment of administrative costs should be based on the basis of an **assumption of full compliance by all entities concerned**. All the assumptions concerning population size (e.g. SMEs), in particular for proposals with long time horizon, should be clearly explained.

Step 9. Assessment of the performance of a 'normally efficient entity'

In order to keep assessment of costs at a reasonable level and ensure comparability, the **assessment will be based on representative types** (typical firms, typical public service, etc.).

The assessment begins with a critical review of available data, identification and removal of obvious outliers (entities whose performance is clearly atypical, i.e. greatly below or above the other performances). In many cases, calculating the median or the average of remaining data might be sufficient.

In addition to the **number of minutes**, the 'normal' level of qualification required by the main actions linked to information obligations and the **'normal' labour cost per hour** including prorated overheads (expenses for premises, telephone, heating, electricity, IT equipment, etc.) will have to be determined.

Step 10. Assessment of the 'business as usual' costs and extrapolation of validated data to EU level

As the aim of the EU standard cost model is to assess the additional burdens originating from EU legislation, one should assess what **part of the costs of administrative obligations would have been borne even in the absence of EU legislation**. For instance, if a company is obliged under an EU Regulation to report on greenhouse gas emissions but would have been obliged to report under national rules on CO_2 emissions only, the burden stemming from EU legislation relates only to reporting on additional greenhouse gases and the 'business as usual' factor amounts to, e.g. 60%.

There is **no need to provide specific estimates for each Member State** or administrative body concerned, unless to do so would be proportionate. In most cases, EU costs will be

estimated by extrapolating available data at national or EU level by considering the number of affected entities across the EU.

Step 11. Report

Estimates of administrative burdens **need to be reported in an impact assessment at least for the preferred option in the obligatory Annex 3 ('who is affected and how')**. If the initiative is likely to generate significant administrative burdens by its nature or such costs are a significant distinguishing factor among the considered options, they should be reported for all the options.

For any initiatives introducing new or removing existing recurring administrative burdens, such costs should also be reported by means of a dedicated online OIOO calculator[844] that supports reporting on the application of the 'one in, one out' approach.

[844] The online OIOO calculator

TOOL #59. COST ESTIMATES AND THE 'ONE IN, ONE OUT' APPROACH

1. INTRODUCTION

The Commission has committed to the 'one in, one out' (OIOO) approach[845]. This means **offsetting new burdens**[846] **resulting from the Commission's proposals**[847] **by reducing existing burdens in the same policy area**. The 'better regulation' Communication of 29 April 2021, COM(2021) 219 final, sets out the main principles of the approach. Burdens shall be identified through cost estimation methods. The present tool gives practical guidance on how to estimate costs and report them for the purpose of OIOO. The tool links closely to tools #56 (*Typology of costs and benefits*), #57 (*Methods to assess costs and benefits*) and #58 (*EU Standard Cost Model*).

2. SCOPE OF COST ESTIMATES

OIOO is based on cost estimates for how proposals affect businesses and citizens. Hence, **impact assessments accompanying proposals will need to consider whether these would have significant**[848] **cost implications for businesses and citizens**[849]. This will need to be done already at the planning stage to identify necessary data collection and analytical work to be undertaken for the impact assessment.

In the Decide entry, services should specify if the case is likely to have significant costs implications (add or remove compliance costs, i.e. administrative and/or adjustment costs). The REFIT box must be checked, if the file is a revision of existing legislation, as all revisions must seek burden reduction and simplification, or justify if this is not the case. For relevant cases, necessary data collection and analytical work will have to be planned and carried out.

For most proposals with significant cost impacts on businesses and citizens, an impact assessment will be produced. This is because significant cost implications would tend to overlap with the 'significant impact' criterion for impact assessments[850]. In these cases, the impact assessment will be the basis for the calculation and presentation of the cost estimates. In cases where the Commission was unable to produce an impact assessment where one should have been prepared, cost estimates will be presented in an analytical SWD, which shall present and justify the cost estimates made[851].

---

845 See Political Guidelines of President von der Leyen, 'better regulation' Communication COM(2021) 219).

846 As defined in Tool #58 (*EU Standard Cost Model)*

847 Legislative proposals and delegated and implementing acts.

848 In line with proportionality, cost estimates are prioritised for cases with significant costs. This means that OIOO does not cover proposals, which entail negligible administrative costs for businesses and citizens. Costs may be negligible either because the proposal does not involve businesses or citizens very much or because ins and outs in a proposal offset each other. 'Significant costs' relates to the overall magnitude of a burden as stemming from the substance of the requirements, the number of affected business/citizens (i.e. the impact on a large number of business, relevance for SMEs and micro businesses, etc.). If in doubt, consult the 'better regulation' functional mailbox.

849 Consequently, OIOO does not concern cost implications for Member States, national and regional authorities or public administrations. Furthermore, OIOO only concern businesses in the EU and EU citizens.

850 See Tool #7 (*What is an impact assessment and when it is necessary*)

851 Where an impact assessment should have been prepared, but it wasn't, estimates of costs and benefits should be presented in analytical reports alongside the proposal or within 3 months, explaining clearly how they have been calculated.

Commission services must identify relevant cases with significant costs implications – starting with the **Commission work programme** and make sure that the required work is undertaken to produce all relevant estimates. The Secretariat-General will help ensure that relevant cases are identified and that the OIOO procedures are followed.

3. THE NATURE OF COMMISSION COST CALCULATIONS

The Commission will not be able to anticipate how Member States intend to transpose, implement, and enforce its proposals, after they have been amended and adopted in the legislative process. For this reason, cost calculations at the level of the Commission legislative proposal do not account for the specific implementation, enforcement modes and efficiency of public authorities in 27 different Member States. Instead, Commission cost calculations are based on working assumptions simplifying the complex reality of the Union and are meant to convey magnitudes of the costs involved to policy makers. For this reason, describing assumptions and uncertainty is important.

Costs can be estimated based on a variety of sources. These may include:

- data obtained from Member States, public authorities, agencies, etc.;
- data from public or targeted consultations, stakeholder workshops, surveys or similar;
- evidence provided by social partners and representatives of economic and professional interests;
- results of SME panels;
- evaluations;
- reports and studies from consultants and academics.

4. APPLICATION OF THE 'ONE IN, ONE OUT' APPROACH

4.1. Compliance costs

Compliance costs are all costs directly linked to the compliance with a law. These consist broadly of adjustment and administrative costs. The OIOO mechanism rests on these two legs.

- **Adjustment costs** refer to incremental costs for investments, equipment, etc. relating to adjusting to the substantive legal requirements of proposals. See Tool #56 (*Typology of costs and benefits*) for a precise definition.
- **Administrative costs** are costs that result of administrative activities performed to comply with administrative obligations included in legal rules. They concern costs for providing information, such as notification of activities, submission of reports, information labelling and certification of products or processes. See Tool #58 (*EU Standard Cost Model*) for precise definitions.

All relevant compliance costs or savings should be estimated, be it adjustment or administrative costs, one-off or recurrent costs, direct or indirect costs, to the extent that this is feasible. Tools #57 (*Methods to assess costs and benefits*) and #58 (*EU Standard Cost Model*) provide all the steps to follow for cost estimations. If it is not possible to provide a quantitative estimate, a qualitative estimate should be made. The Regulatory Scrutiny Board will review costs estimates as part of the regular scrutiny of impact assessments.

Adjustment costs and administrative costs will be treated in a different way within the OIOO approach, as follows.

4.2. Adjustment costs

DGs shall estimate adjustment costs for business and citizens in impact assessments, to the extent that this is feasible and proportionate.

Adjustment cost are the incremental costs – other than fees and administrative costs – to the target group. They cover elements such as direct labor costs, overheads, equipment costs, material costs, cost of external services and implementation costs[852].

Box 1: Steps for estimating adjustment costs

1) Identify substantive requirements that legislation (in line with the preferred option) will impose on citizens and/or business[853];
2) Estimate the number of businesses and citizens who will be affected by the proposal.
3) Estimate the adjustment net costs for a 'normally efficient business' or citizen;
4) Estimate the 'business as usual' factor[854];
5) Consider special groups (SMEs, micro businesses, etc.);
6) Estimate adjustment costs for every substantive requirement;
7) Estimate whether adjustment costs are likely to change over time (i.e. correct for digitalisation);
8) Sum up and extrapolate[855], if necessary.

Adjustment costs shall be presented transparently and systematically in impact assessments when discussing the preferred option. A specific section will illustrate the application of the OIOO approach for all relevant initiatives in the following way:

- adjustment cost estimates for business and citizens will be illustrated for each category, where possible;
- benefits will be quantified to the extent possible, as an order of their magnitude (or at least with a solid qualitative analysis);
- sources of estimates will be referenced, and data limitations shown transparently;
- if models are applied, they will be referenced in MIDAS;
- where possible, existing or included compensatory measures[856] for business will be identified.

[852] See Tool #56 (*Typology of costs and benefits*) for more details.
[853] If the impact assessment has no preferred option, the cost estimates will reflect the choice of the actual proposal. In this case, cost estimates can be presented in the explanatory memorandum.
[854] 'Business as usual' costs are those that the business would have incurred even in the absence of the proposal. For instance, a proposal may lead to changes, which have already been introduced in national legislation in some Member States or it may already be a business practice in certain sectors. Certain costs may turn into 'business as usual' costs over time as they are integrated in standard business practices.
[855] In some cases, individual company data or data from a few Member States will need to be extrapolated to the EU level.

Adjustment costs and benefits will be specified in the obligatory summary table of the Annex 3. Under the preferred option, relevant initiatives will summarise:

- adjustment costs estimates (as explained in box 1) and benefits of the proposed option[857];
- administrative costs that will be offset.

Adjustment costs in the short term can lead to business opportunities in the longer term. For example, the twin transition might entail short-term costs and financing that, in the medium term, will make Europe more resilient to the benefit of the wellbeing of citizens and firms, by allowing business to have first-mover advantages, acquire new markets and introduce more efficient business models. Some of these advantages cannot be monetised and go beyond GDP[858], to increase the overall resilience of the system.

Adjustment costs will be **compensated** to the greatest possible extent to cushion the short-term effects of these measures and allow businesses to overcome any adjustment and other costs in the short term. An example of such broad compensations are initiatives such as the Next Generation EU[859].

Companies may also benefit generally from the support provided to other sectors, for instance, education and research, infrastructure, etc. including that directed to Member States, public entities, etc.

These compensatory measures will help mitigate the costs of adjustment. They will provide broad or targeted relief, which will support business facing these costs. These compensatory measures, where the case, will be presented in the relevant impact assessments to the extent possible. Any further follow up compensatory measures will be presented in the Annual Burden Survey.

4.3. Administrative costs

Administrative costs, when significant, will be **offset**[860]. To do so, they need to be identified, estimated, reported, and made comparable (monetised) so that they can be aggregated. While the standard cost model is useful (see Tool # 58 (*EU Standard Cost Model*)), and services should strive to apply it as much as possible, it is not obligatory to apply all its dimensions, given that the information required for it may not be available (such as estimated time used, or the number of businesses affected).

---

[856] Impact assessments shall present mitigation measures taken in context of the proposal and/or refer to compensatory measures already in place.

[857] Compensatory measures, where the case, will be presented in the relevant impact assessments to the extent possible.

[858] This is in line with the Commission's ambition, stated in COM(2020) 493 final, to find new metrics to measure progress beyond GDP, whereby resilience dashboards are being developed.

[859] Recovery plan for Europe | European Commission (europa.eu)

[860] In many cases businesses would have to bear costs even if there was no obligation at the EU level in a given policy area (the 'business as usual' costs). The offsetting is applied only to the remaining administrative burden defined as administrative costs minus 'business as usual' costs, see Tool #58 (*The EU Standard Cost Model*). The OIOO calculator will help services calculate the right estimate based on information about administrative costs, etc.

Both administrative costs for citizens and business will be reported. Cost for citizens will be treated on a par with costs for business.

Several issues should be considered when undertaking estimates for administrative costs:

- **Consideration of costs in European Commission proposals only:** OIOO only applies to cost implications originating from Commission proposals. Hence, costs imposed by other parties – for instance added by co-legislators or by Member States and local, regional authorities – are not included[861].
- **'Business as usual' costs are excluded[862]:** Cost estimates should only cover the impact of new regulatory requirements. For instance, for some Member States, a certain administrative requirement may already be in place. Hence, to the extent that new elements overlap with existing processes, the 'business as usual' costs should be subtracted in cost estimates for OIOO. This goes also for revisions, where cost estimates should only account for new costs added or subtracted, not costs incurred by existing legislation. Information available as 'business as usual' information may however vary from one Member State to the other and between sectors as well. This should be accounted for to the extent possible. See Tool #58 (*EU Standard Cost Model*) for more precise methods to account for 'business as usual' costs. When introducing a new requirement, administrative cost may be significant, but, over time, requirements may become normalised in the business context – integrated in the standard business processes and accounted for in the general framework. Thus, costs from new requirements may become 'business as usual' costs over time.
- **Account for single market simplification:** In cases, which involve a certain harmonisation of requirements across EU Member States, the resulting simplification of the regulatory framework in the single market should be considered. For instance, when one common EU standard replaces 27 different standards, businesses active cross border in the internal market will experience a simplification in the regulatory environment that needs to be considered. Plausible assumptions may have to be made, for instance based on existing performance of Member States or lessons learned in evaluations. See dedicated tool on the internal market (Tool #25).
- **Cost implications in evaluations and fitness checks:** Cost implications (in particular the 'outs') are not always linked to important new proposals. They may occur, for example, when digitalisation projects streamline existing reporting and monitoring processes; or when initiatives bring important legal clarity to a policy area, where practices have previously been diverging and uncertainly prevailed. Evaluations and fitness checks can also lead to corrections in previous cost estimates by providing real evidence of how requirements are handled in reality. This may lead to adjustments. Such cost implications may be documented in evaluations or fitness checks[863].

4.4. Reporting administrative costs

For administrative costs, the OIOO online calculator should be used. The Commission has developed a dedicated calculator for reporting administrative costs for the purpose of

[861] Costs imposed through proposals transposing international agreements already endorsed by Member States are also exempt from OIOO. Legal proposals reflecting European Citizens' Initiatives are exempt.
[862] See Tool #58 (*EU Standard Cost Model*) for further info about 'business as usual' costs.
[863] Such cost implications will have to be linked to Commission initiatives. There will obviously be a certain time lag in the recording of such estimates.

OIOO. The calculator is based on the EU standard costs model but allows aggregation by policy area and greater flexibility in reporting data. Nonetheless, it will remain important to be transparent in the impact assessment report about how the estimates have been produced and on which basis they rely.

Services shall make sure that every proposal with significant administrative cost implications for business or citizens is recorded in the OIOO online calculator[864]. Impact assessments are the standard basis for this recording.

The calculator provides a flexible module for recording cost estimates (ins and outs) for offsetting under OIOO. Cost estimates shall be recorded in the tool at the latest at the time of adoption of the Commission proposal[865], although it is recommended to do this at the time of submission of the documents to the Regulatory Scrutiny Board.

If you have practical difficulties in applying it, please contact SG-BETTER-REGULATIONS@ec.europa.eu. If you have IT related questions to the calculator, please contact the IT helpdesk.

Box 3: Requirements for reporting OIOO administrative costs

Cost estimates for OIOO to be reported in the calculator will have to be calculated as follows:

1. Entries in the calculator shall result in one single administrative burden (administrative costs minus business-as-usual costs) estimate per entry (impact assessment/ proposal/initiative). This can be positive (ins) or negative (outs). One single amount per act[866] or per initiative per year (in EUR). While there can be different underlying elements (ins and outs), these will have to be summed up to provide one single number per initiative.
2. Cost estimates shall be specified in absolute amounts (EUR). Thus, it is not possible to specify costs per procedure, per unit or as a share of turnover or investments or similar.
3. Costs can be estimated and reported in the impact assessment and elsewhere as a range with a minimum and maximum, reflecting the level of uncertainty of the estimate. However, for the purpose of OIOO, a single value (for example the average/median value) will have to be reported. The level of uncertainty can be indicated in the comments field.
4. Costs reported from proposals with joint DG responsibility can be assigned to the DGs proportionately.

[864] Access to the tool is restricted to the 'better regulation' functions in the DGs. For receiving access to the tool, please contact SG BETTER REGULATIONS.

[865] In cases where an impact assessment was not carried out due to urgency, a SWD will be published maximum three months after adoption (see COM (2021) 219). In such cases, the deadline for recording cost estimates in the online OIOO calculator will naturally follow the SWD.

[866] For cases where the impact assessment results in more than one legislative initiative, the administrative costs should be introduced in the impact assessment, i.e. per initiative, not per each legislative proposal. These may include directives/regulations as well as delegated and implementing acts.

4.5. Offsetting of administrative burdens

The offsetting of administrative burdens will take place through an **offsetting hierarchy**:

1) First, burdens will have to be offset within the remit of each DG.
2) If this is not possible, burdens will be offset within the REFIT area[867] directly affected by the proposal.
3) If this is not possible (e.g. in the case of a cross-sectoral initiatives or where the Directorate-General does not produce sufficient legislation or the existing legislation in the policy domain is recent), the costs will be offset within the broader policy areas, i.e. the Commission work program's headline ambitions. The six headline ambitions of the European Commission define the policy areas:

   - A European green deal;
   - A Europe fit for the digital age;
   - An economy that works for people;
   - A stronger Europe in the world;
   - Promoting our European way of life;
   - A new push for European democracy.

4) If the proposed legislation is deemed to be necessary, but it is not possible to find an 'out' in the same area, the Commission can, in exceptional cases, decide to take the 'out' from a different policy area, thus 'trading' can take place **between policy areas**.
5) If an 'out' cannot be identified in the same year, costs will be reported to the next year. Hence, OIOO accounts will not balance every year, but should balance **over time**.

Furthermore, the Commission can – in exceptional cases – decide to exempt an initiative from OIOO and from offsetting. This may apply to regulations in emerging areas, where it is necessary to fill a regulatory gap. Services should signal the request for exempting an initiative from OIOO and explain the reasons. The exemption is granted by the College following a presentation by the Vice-President responsible for 'better regulation', and the explanation will be provided in the explanatory memorandum of the proposal.

Services who wish to exempt a proposal from the 'one in, one out' approach should send a reasoned request to the functional mailbox: SG-BETTER-REGULATION-EXCEPTIONS@ec.europa.eu.

[867] i.e. agriculture and rural development; climate action; communication networks, content and technology; competition; employment, social affairs, inclusion and education; energy; environment; financial stability, financial services and capital markets union; health and food safety; internal market, industry, entrepreneurship, SMEs and statistics; justice, consumers and gender equality; maritime and fisheries policy; migration and home affairs; mobility and transport; regional and urban policy; research and innovation; taxation, customs union and anti-fraud; trade and external action.

5. FINAL COMMENTS

While DGs shall prepare the cost estimates and enter them in the dedicated OIOO online calculator, Secretariat-General will aggregate estimates in view of offsetting and reporting for the Annual Burden Survey. Specific impact assessment, SWDs and explanatory memoranda shall refer to offsetting measures only when these are inherent to the proposal. The offsetting will be reported in the Annual Burden Survey, based on the offsetting hierarchy indicated above.

6. REFERENCES

- 'Better regulation' Communication (COM(2021) 219 final)
- Dedicated OIOO online calculator
- 'Better regulation' toolbox, notably Tools # 56 (*Typology of costs and benefits*), #57 (*Methods to assess costs and benefits*) and #58 (*EU Standard Cost Model*).
- OECD: One-In, X-Out: regulatory offsetting in selected OECD countries (OECD regulatory policy working paper). https://doi.org/10.1787/67d71764-en
- OECD (2014), OECD Regulatory Compliance Cost Assessment Guidance, OECD Publishing. http://dx.doi.org/10.1787/9789264209657-e
- European-Commission-Admin-Burdens-Cost-Model-Manual-2006.pdf (regulatoryreform.com)

TOOL #60. BASELINES

1. WHAT ARE BASELINES?

The 'better regulation' guidelines state that the design of possible policy options should always consider the option of changing nothing and use this as the **benchmark against which the policy option(s) should be compared**. This benchmark is usually referred to as baseline[868] (scenario).

As described in Tool #11 (*Format of the impact assessment report*), a baseline is a 'no-policy-change' scenario which **includes all relevant national, EU-level and global policies and measures which are assumed to continue to be in force**. In addition, (relevant) Commission proposals even if not yet adopted by the co-legislator can also be included[869].

In general, one should define the baseline such that it allows to assess the environmental, economic and social effects of the policy initiative. For the sake of coherence, where two or more initiatives are being developed simultaneously, or presented together (e.g. as a 'package') each impact assessment report should use the same baseline. It may be useful to consider alternative baselines to demonstrate the impacts of other related initiative(s) or proposals.

Baselines are not only relevant in impact assessments (as defined in Tool #11), but also in in evaluations (Tool #46 (*Designing the evaluation*)). Due to the different nature of the evaluations, their baselines can be different from the general approach for impact assessments. The evaluation will assess the intervention against this baseline (or point of comparison). This can be for example the situation before the intervention started (T0), the dynamic no-change scenario in the impact assessment (baseline), or the situation expected to be achieved at that point in time (preferred option). Different points of comparison may help to assess different evaluation criteria and the evaluation should clearly explain the choices made.

Both for evaluations and for impact assessments, the baseline should develop an appropriate time horizon to allow for the expected impacts to be realised. It will also be relevant to include expected socio-economic developments (aging, GDP growth, etc.) as well as important technological/societal developments, including foresight elements such as megatrends (Tool #20 (*Strategic foresight for impact assessments and evaluations*)). Baselines should be quantified as far as possible, although in many cases baselines of impact assessments and evaluations are of qualitative nature.

Much of the remainder of this tool will focus on the quantification of baselines, alternated with recommendations that are valid for both quantified and qualitative baselines.

2. QUANTIFICATION OF BASELINES

In the European Commission, the reference scenarios for Energy, Transport and Green House Gas Emissions or the Agricultural Outlook are examples of detailed, quantified and

[868] Alternatively, the term 'reference scenario' is used instead of 'baseline scenario', but the two notions usually carry the same meaning.

[869] Where it is clear that a subsequent legislative procedure will deliver a substantially different outcome to the Commission's original proposal, this outcome should also be reflected in the baseline.

comprehensive scenarios, which are produced regularly and involve a large number of variables and assumptions.

A baseline scenario is a projection based on a number of baseline assumptions and status quo policies as input(s). Broadly speaking, baseline assumptions are all (implicit and explicit) assumptions underlying a baseline scenario. These baseline assumptions could come from different methodologies such as a forecasting model; or alternatively be based on expert judgement. The values of these baseline assumptions may (or may not) change over the time horizon of the assessment. Relevant assumptions for most baselines include expected socio-economic developments (aging, GDP growth, etc.), technological/societal developments or megatrends. As such, baselines can be one of the channels to bring foresight (Tool #20) into the 'better regulation' framework.

2.1. Why do baseline assumptions matter?

The impact of a policy option typically is compared with a clearly defined baseline. If the baseline is ill-defined or poorly specified, then so will be the estimated effects of the policy.

Figure 1: Baselines and policy effects

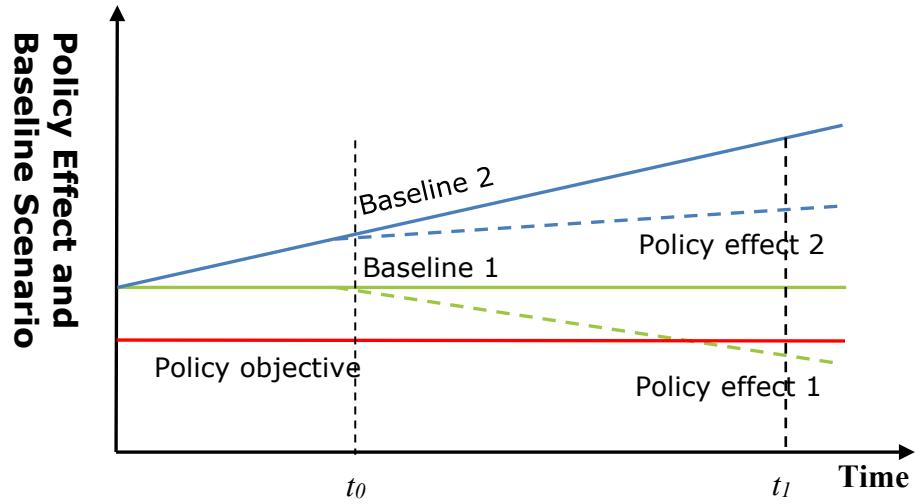


Figure 1 illustrates the importance of a well-defined baseline. Consider a policy action that seeks to curb CO_2 emissions from road transport in absolute terms – the policy objective – by fostering the fuel efficiency of cars through restrictions on fleet consumption[870]. If the analyst assumes (in Baseline 1) that the total mileage (number of vehicles times kilometres driven) of cars remains largely constant, then a 10% improvement of fuel efficiency should reduce emissions by approximately 10% in the course of time as more efficient cars are put into service. If, on the other hand, total mileage is assumed to increase by 20% over time (in Baseline 2), say because there are more vehicles on the road, then a 10% efficiency improvement will not reduce emissions as required and additional measured may be needed.

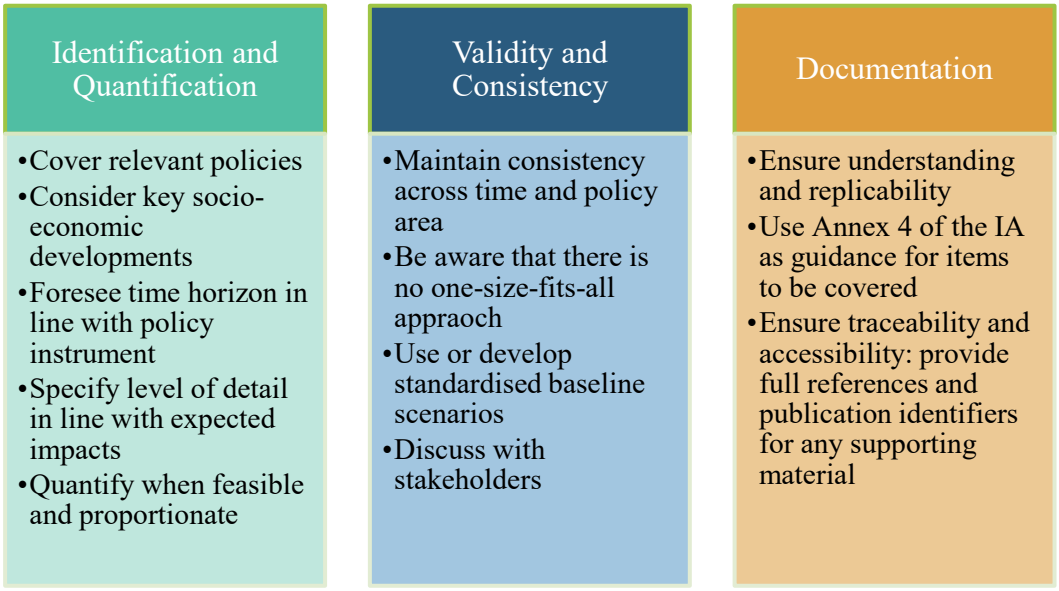
Since baseline assumptions are likely to evolve over time (for instance because of other ongoing policies or other exogenous changes), the policy effects in t_1 should ideally not be compared with the current situation (i.e. the situation in year t_0 or the *status quo*), but with the baseline in time horizon t_1.

[870] The example provided is illustrative and only serves to exemplify a complex situation.

Practical implementation steps

The identification of baseline assumptions and, subsequently, of a baseline scenario requires three key steps: Identification and quantification, Validity and Consistency checks, and Documentation (Figure 2).

Figure 2: Steps to be taken



2.2. Identification: what is to be considered?

Baseline assumptions should **include all relevant national, EU-level and global policies and measures,** which are assumed to be in force or to be implemented by the time the new policy proposal will take effect. In the above example in Figure 1, relevant policies include all policies other than the one to be assessed, which have an impact on the future mileage of cars. These may be closely related policies, such as excise taxes on fuel, but also policies in more remote fields such as investment in transport infrastructure. Where it is clear that a subsequent legislative procedure will deliver a substantially different outcome to the Commission's original proposal, this outcome should also be reflected in the baseline.

Where the impact assessment concerns regulatory initiatives based on a legal obligation for the Commission to act (e.g. through delegated or implementing acts), **the baseline should be construed as a 'no action' reference scenario** which, serves primarily as a benchmark for comparing options.

Expected socio-economic developments and trends (aging, GDP growth, etc.) as well as **important technological/societal developments** inside and outside the EU, such as the pervasive nature of the internet, social media, and emerging technologies should be considered, provided they are deemed relevant for the policy field of the impact assessment, the baseline scenario, and the expected effects of the policy. Using again the above example, both growing GDP and shrinking household size may lead to an increase in the number of cars and thus change the baseline while increasing fuel prices may reduce the use of cars.

A complicating factor is that **the policy or legislation itself might envisage already an end date ('sunset clause')** and that a positive decision of the Commission and Legislator will be necessary to put in place a new policy regime. In such cases, it can be difficult to decide on the appropriate baseline. Two options are possible:

- Option 1: include the 'sunset clause' in the baseline if a comprehensive evaluation concludes that the policy is ineffective. Policy options would then include establishing a new action and the impacts would be measured against a no-policy baseline. This approach should, however, be avoided if there are clear political commitments to continue the policy in some form or another for reasons other than its effectiveness.
- Option 2: include a continuation of the current policy approach in the baseline even if it formally comes to an end, where, for example, a comprehensive evaluation concludes that the policy is effective. Given that the College or Legislator could (theoretically) decide not to propose or enact legislation, this approach should usually be accompanied by a policy option, which would explicitly repeal the current policy and would demonstrate the cost of the Union not acting ('the cost of non-Europe').

The most appropriate option has to be decided on a case-by-case basis considering the degree of political commitment to a continuation of the current policy and the results of evaluations and fitness checks.

Baseline assumptions should be determined for an **appropriate time horizon**. Its length depends on the likely lifetime of any individual option and on the expected timespan for impacts to be realised (which may occur long after the policy has ended).

Also the **level of detail and precision** should be chosen with care in order to avoid information overload while capturing the effects that are to be examined. In model-based impact assessments, the baseline assumptions of the impact assessment correspond in principle to those of the underlying model(s) and should therefore cover all relevant assumptions underpinning the model itself, its application and implementation.

Baseline assumptions and the resulting scenario(s) should, whenever and wherever possible**, be quantified**. In doing so, analysts should draw upon the best available evidence, expert judgement and scientific knowledge, and being aware of plausible surprise events or other kinds of discontinuities. Such events, for instance a political conflict leading to a sharp increase in fuel prices, may have a significant impact on the environment, in which the policy measure is to be implemented, but do not result from a projection of past trends and may therefore be difficult to quantify. A false degree of precision should be avoided, meaning that it is important to be transparent with the limitations and error margins of the quantification values.

2.3. Consistency of baseline assumptions and scenarios

A set of baseline assumptions is consistent (with another set) if there are no logical and factual contradictions including significant differences within or between both sets. However, a baseline is also context-dependent and there may be differences across baselines for various reasons:

- **Same model – different policy areas**: when a model is used for assessing different policy fields, different (ideally, non-contradicting) baselines can be used.

- **Same area of analysis – similar models**: if similar models are used in the same policy field(s), their baseline may differ somewhat due to the different model structure and implementation of the (same) baseline assumptions (e.g. E3ME[871] vs. JRC-GEM-E3[872]).
- **Same** impact assessment **– different models with shared assumptions**: if the impact assessment makes use of various models whose underlying assumptions overlap to some extent, i.e. which share some of the variables of the other models used in this impact assessment (e.g. GEM-E3 vs. POLES[873]).
- **Different** impact assessment **with different models with overlapping assumptions**: when impact assessments for different policy fields use different models with nevertheless significantly overlapping assumptions and variables such as data on GDP or population trends.
- **Same model – different points in time**: if the same model is used repeatedly over time in order to analyse similar policy actions, each time with a different set of assumptions.

Standardised baseline scenarios

Contradictions between baseline assumptions can undermine the credibility of impact assessments, and if they occur, should be explicitly recognised and fully justified. A comprehensive set of assumptions in a numerical context (e.g. the EU Reference Scenarios) help identify possible inconsistencies.

While avoiding contradictions can be regarded as a general principle of good governance, one should strive for consistency to facilitate comparability of impact assessments across time, policy area, or institution. From this, it follows that wherever possible, standardised baseline scenarios could be developed and used for various related policy fields.

The EU Reference Scenario for energy, transport and climate policies is a prime example. It is updated regularly and built upon a modelling framework in which simulation models across various policy fields consistently interact with each other. While the Reference Scenario has been developed for specific policy fields, it is also used in others. Consider therefore whether (parts of) the Reference Scenario is suitable for the impact assessment under preparation.

The yearly EU Agricultural Outlook is another example, based on a consistent set of assumptions regarding macroeconomic conditions, the agricultural and trade policy environment and market developments.

A third example are the EU and Global macroeconomic baselines which integrate macroeconomic projections and sectoral information in a consistent Input-Output framework.

Sharing baseline scenarios ensure consistency, transparency, and reduce the costs of carrying out impact assessments. At the same time, models can interact with each other through shared scenarios and assumptions whenever that is required for a specific impact assessment.

[871] https://www.e3me.com/
[872] https://ec.europa.eu/jrc/en/gem-e3
[873] https://ec.europa.eu/jrc/en/poles

The preparation of reference scenarios requires close cooperation between modellers and policymakers as well as other relevant stakeholders such as national experts, market experts, and experts from private companies and international organisations.

Justified differences between baselines

The above arguments imply by way of inversion that baseline assumptions can be different in **some situations**. The most obvious case concerns impact assessments which are carried out at different points in time as this may imply that:

- data take on (a) different value(s) at different points in time,
- new evidence becomes available and/or previous evidence becomes obsolete,
- data collection or sampling methods (including measurement methods) are newly developed that allow a modified parametrisation of the model.

Other reasons for justified differences include:

- different sectoral, temporal or spatial granularities of data, e.g. because more data are collected allowing for further differentiation among sampled entities,
- different data requirements for the models used along the above dimensions of granularity, e.g. because some models allow for greater detail than others.

Before reusing a baseline scenario, analysts should therefore check whether the underlying assumptions are still fully justified or whether an update is warranted.

2.4. Documentation

All assumptions that have been made in the course of the analysis should be **documented in Annex 4** to the impact assessment (see Tool #11 (*Format of the impact assessment report*))[874]. This applies in particular to all numerical assumptions and chosen model parameters, if applicable, and other relevant numerical values. Where the analysis makes use of models or methodologies developed by third parties, only the idiosyncratic assumptions should be identified and documented. For all other assumptions, it is sufficient to provide references to manuals or other guidance that allow third parties to obtain information on these assumptions, as long as these references reflect the assumptions at the time of the impact assessment[875].

3. ADDITIONAL INFORMATION

- EU Reference Scenarios for energy, transport and climate policies
- EU Agricultural Outlook
- EU and Global macroeconomic baselines for policy assessments

[874] If your assumptions are equal to those used in a past impact assessment, please document them in Annex 4 anyway, to ensure maximum transparency. This is particularly true if the assumptions are based on those of a past impact assessment or reference scenario but have been partially updated to match the evolving policy landscape.

[875] Note that especially manuals have the tendency to evolve over time. To ensure that all relevant information to understand the impact assessment remains traceable also considerable time after the impact assessment has been carried out, the references to manuals should include versioning information, and each version must remain traceable over time.

TOOL #61. SIMULATION MODELS

1. MAIN FEATURES

This tool gives an introduction to some practical considerations, if one choses to use models for a policy analysis. **Models** are stylized representations of the real world that are used to make projections or to assess the behaviour of a system under specific (policy) assumptions. Models can provide support to policy makers throughout the policy cycle and across a wide range of policy areas. While all models are simplifications of reality, good models can provide useful insights and understanding if appropriately used. A good model selection and use is crucial to deliver high quality, policy relevant results. Simulation models can assess a wide array of **economic, social and environmental impacts**. A detailed description of all model types is well beyond the scope of this tool; some examples of commonly used model types for Commission impact assessments are listed in Box 1.

Box 1. Models used in impact assessment

Economic models used in impact assessments include g*eneral equilibrium models*, that allow for consistent comparative analysis by ensuring that the economic system and individual markets remain in equilibrium in the long term; *econometric models*, estimated using historical time-series data, to capture medium/long-term effects from shocks and for forecasting; *partial equilibrium models*, used in the detailed analysis of one or more specific economic sectors over the short/medium/long term; and *micro-simulation models*, typically used for analyses at a detailed disaggregated level over the short term focusing on individuals, households or firms.

Environmental models can represent dynamics and impacts on resources such as water, air, land and on biological processes or assets. **Energy models** simulate energy systems and markets, while **transport models** allow the analysis of transport activities, related energy consumption and air emissions. Models are also used in a variety of other fields, such as to simulate **land use** and **population dynamics**.

By combining several models together, **integrated modelling approaches** allow assessing impacts in several policy areas. Ensuring coherence is of major importance when developing and building linkages between the models.

Ex-ante policy analysis and impact assessments can look at sustainable development through the overarching **sustainable development goals (SDGs)** framework, and models are a fundamental tool to understand the complexity of sustainability and development challenges, which often lie in multidisciplinary domains. Furthermore, achieving sustainable development requires a good understanding of the interlinkages (synergies and trade-offs) among the goals, and modelling can provide a concrete contribution to this. The JRC SDG Platform[876] maps models with the SDGs framework at goal and target level, and links them with the EUROSTAT and UN SDGs indicators. This is a useful tool to support policymakers in identifying the appropriate model(s) for the assessment of their specific policy options.

In **impact assessments (IA)**, models are extensively used to assess the environmental, economic, and social impacts of policies. They are used for the problem definition, to contribute to the construction of baselines, and for the evaluation of existing policies. 16% of

[876] https://knowsdgs.jrc.ec.europa.eu/. See also Tool #19 (*Sustainable development goals*).

all IAs carried out by the European Commission in the years 2003-2018 are supported by models, growing to 25-30% of IAs from 2015 onwards. Around 120 models have been used to this end (Acs et al., 2019). Policy areas characterised by frequent use of models in support of policy making in the EU include climate and environment, economics, transport, energy, trade, agriculture, and the complex interaction between sustainability and development (see Box 1).

Models can be used alone or in **combination**. Complementarities between different models should also be explored, as model types could offer a more accurate representation of some impacts in respect to others. Models can consistently interact with each other through shared scenarios and assumptions. For example, the development of EU Reference Scenarios combines models for the assessment of the impacts on economy, energy, transport, land use, agriculture, forestry and air quality[877]. Figure 1 presents models used to inform the European Commission's 'A Clean Planet for all' (COM (2018) 773).

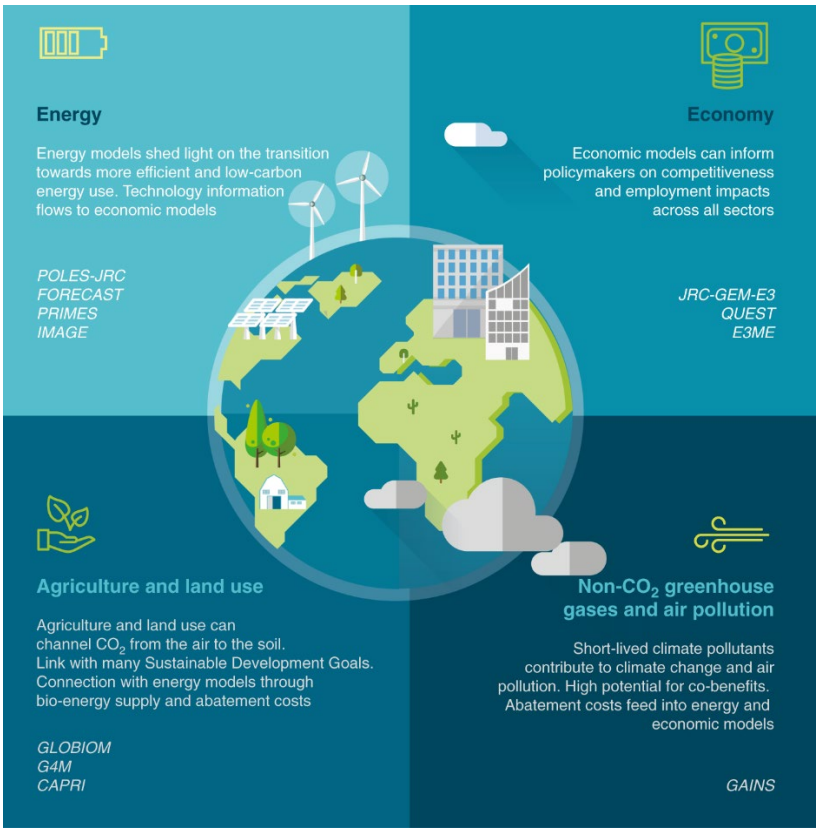


Figure 1: The modelling toolbox (model names are at the bottom of each box) to inform the European Commission' 'A Clean Planet for all - A European strategic long-term vision for a prosperous, modern, competitive and climate neutral economy' [COM (2018) 773][878]

Modelling is a complex and technical activity for which **specific expertise** is needed. Models can have different spatial coverage and detail, and can represent short-, medium- or long-term effects. Input data requirements usually are demanding, and data quality is fundamental for the quality of model output. Specific software is needed for handling the input data, running model simulations, and interpreting and presenting the results.

877 https://ec.europa.eu/energy/data-analysis/energy-modelling/eu-reference-scenario-2016.

878 Weitzel, M., Vandyck, T., Keramidas, K. et al. Model-based assessments for long-term climate strategies. Nature Climate Change 9, 345–347 (2019).

Models are developed and run **within the Commission** (JRC and policy DGs), as well as by **external organisations** on behalf of the Commission, such as research institutes, universities, and private consultancies. In general, the development of a new model is a complex, resource- and time-intense process. Sufficient and long-term resources are needed to ensure model maintenance over time. Developing a model and preparing scenarios in light of a new policy initiative therefore needs to be anticipated well in advance. JRC has multi-annual administrative arrangements with many DGs, many of which include modelling.

Model development and use entail **close collaboration over time between modellers and policymakers**. On the one hand, models should be developed, managed and maintained by appropriately skilled and experienced staff. On the other, decision makers should understand the general model functioning, limitations, risks, major assumptions and outputs. This can be ensured by close communication with modellers on how a model works and the strengths and limitations of a specific approach. This mutual interaction is key to build trust and ensure a high-quality outcome. Trust in the models in use by the European Commission can be further strengthened through transparency and by engaging Member States' experts in the validation of the data, the modelling results and the baseline scenarios.

2. TRANSPARENCY AND QUALITY ASSURANCE

Models used for policy support should be sound from a scientific point of view and transparent from the perspective of stakeholders. The higher the **quality** and thus reputation of a model, the more convincing the modelling results for policy support. The process to achieve these requires time and resources, typically covering various impact assessment cycles, and is the result of a number of actions. These should be proportionate to the impact and complexity of the model and be an integral part of the work both when models are run by Commission services and by external contractors, through appropriate terms of reference.

Transparency of the model and its use includes the documentation and sharing of model assumptions, architecture, code, data and results (see Tool #4 (*Evidence-informed policymaking*)). It enables scrutiny and replication of results by fellow scientists, reuse of the model over time, as well as better understanding, trust, and acceptance by a wider audience. Documenting models and their use helps policymakers to choose an adequate model when preparing a new impact assessment; the co-legislators to better interpret modelling results; and the general public to get a better picture of the evidence that underpins policymaking. The Modelling Inventory and Knowledge Management System **MIDAS** (see section 3 below), details transparency requirements for model use in impact assessments.

Model **quality assurance** typically includes testing and checking consistency before a new model version is released; if applicable, validation that a model can reproduce historical/statistical data; critical assessment of the model assumptions, in collaboration with the relevant experts and stakeholders, to determine whether they are realistic and relevant to the problem at hand. In addition, as for all impact assessment methods, communicating and understanding **uncertainty in model outputs** is vital. To the extent possible, uncertainty should be quantified and accounted for. This is particularly relevant in an impact assessment as this could change the ranking and conclusions about the policy options (see Tool #62 (*Multi-criteria decision analysis*)). Regular **internal and external review** and discussion in the scientific community (in particular, peer-reviewed publications) are key for the trust in the process and acceptance of the results. **Coherence** could be ensured through a shared baseline, the reference scenario that provides a base for comparison (see Tool #60 (*Baselines*), and Figure 1).

- **Practical steps**

Using models for impact assessment requires close collaboration between modellers and policy officers. As the modelling exercise might require up to several months, it is important to plan well and to perform all preliminary steps in due time to ensure that high quality results can be included in the impact assessment report. Box 2 lists a number of steps that can help informing the impact assessment process with evidence from the modelling exercises.

Box 2: Steps of model use for impact assessment

1 - Assess whether modelling is relevant for your case

A. Look at available evidence, including whether related impact assessments, evaluations, or studies by your or other DGs make use of models
B. Consult the MIDAS modeling inventory
C. Identify which model(s) can simulate the relevant policy impacts

2 - Select which model(s) to use

A. Carefully consider, also by consulting experts and stakeholders:
   i. relevance and appropriateness of model assumptions for your specific case
   ii. model quality
   iii. model transparency
B. Assess if models are available within the Commission or resort to external contractors
C. In case you consider developing a new model, a long-term vision is necessary. Carefully assess time and resource implications.

3 - Design and plan the modelling exercise

A. Define relevant model assumptions
B. Define the baseline consistently with existing reference scenarios (Tool #60 (*Baselines*))
C. Define the policy options scenarios
D. Define and plan over time all steps of the modelling exercise (e.g. data collection and update, model developments, model calibration, model simulation, results' validation)
E. Ensure coordination if more than one model are to be used
F. Check how to document models and their use (see section 3 further below on the Modelling Inventory MIDAS)

4 – Run the model and validate model output

A. Run the model (by Commission Services or external contractors)
B. Account for uncertainty in model results (Tool #65 (*Uncertainty and sensitivity analysis*))
C. Analyze and discuss model results
D. Validate results together with relevant experts and stakeholders

5 - Communicate model results

A. Present results, clearly explaining assumptions and limitations
B. Considering publishing a specific report describing the modelling exercise, to be referenced in the IA report
C. Encourage the experts to publish (at the adequate point in time) a peer-reviewed

	article on the modelling exercise
D.	Ensure permanent traceability of documentation for citing the modelling exercise

- **The Modelling Inventory of the European Commission (MIDAS)**

Coherently documenting these models and their use is a fundamental step for transparent and evidence-based policymaking. All models that make a substantial contribution to the assessment of policy options in an impact assessment have to be described in the **Modelling Inventory and Knowledge Management System of the European Commission (MIDAS)**. **MIDAS** contains the descriptions of models in use by the Commission in support of the policy cycle and is open to the public. This gives access to the descriptions of the models and related contributions to Commission impact assessments. The information available in MIDAS can conveniently be used to complete Annex 4 of the impact assessment reports (see Tool #11 (*Format of the impact assessment report*)) Therefore **inserting information in MIDAS at an early stage of the IA is highly recommended**.

Box 3. Practical actions to document models used for the policy process

Planning & validation phase

You can check in MIDAS if and how modelling was used in previous impact assessments in your field. This can help to assess if the use of models is relevant for your case.

Getting in touch at an early stage of the IA process with the MIDAS team would allow publishing the information in MIDAS timely, to effectively support the audiences of the IA report. In exceptional cases (such as legal barriers), you can decide who can access the information (Commission only, European Parliament and Council, public).

Policy preparation

If you intend to use models in IAs **for the assessment of policy options**, then the models and their contribution **should be described in MIDAS.** In all other cases (support to problem definition, baselines, evaluation of an existing policy), their insertion in MIDAS can be considered wherever relevant and feasible[879].

The **policy DG** leading the IAs is **responsible for the content** to be shared in the public version of MIDAS. By **informing the MIDAS support team at an early stage about intended model use** (see Tool #8 (*What steps should be followed for an impact assessment*)), **you can also get specific support throughout the IA process**: to help you describing the model or to formulate call for tenders or contracts, which ensure that your contractor provides the required information and agrees to its publication in MIDAS under EU copyright.

Concretely, MIDAS describes three different main elements:

1. The Model

The model description includes, according to a predefined structure: **general information** about the model; **details on model structure**; information on **model quality and**

[879] This is the case in particular when models are used for the construction of a consolidated baseline (like in case of the EU Reference Scenario for Energy, transport and GHG emissions Trends to 2050, doi: 10.2833/001137)

transparency and related documentation.

2. The Model contribution to the impact assessment

This refers to how model results contributed to the IA. Concretely, for the assessment of policy options you need to indicate the specific impacts the model helped to assess.

3. Additional information on the full modelling exercise

The modelling exercise refers to comprehensive information on the specific model configuration, any input data and sources, other models involved, as well as information on quality assurance measures. This information is usually either described in a separate annex of the IA report, and/or in a separate study or article referenced in the IA. In MIDAS, **you only need to state who ran the model** for the IA and provide the **reference** to the modelling exercise.

Contact of the MIDAS support team : EU-MIDAS@ec.europa.eu

https://ec.europa.eu/knowledge4policy/modelling/topic/corporate-modelling-inventory-knowledge-management_en

3. ADDITIONAL INFORMATION

- Acs, S., Ostlaender, N., Listorti, G., Hradec, J., Hardy, M., Smits, P., Hordijk, L., *Modelling for EU Policy support: Impact Assessments*, EUR 29832 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-09671-9, doi:10.2760/748720, JRC117250.
- Ostlaender, N., Acs, S., Listorti, G., Hardy, M., Ghirimoldi, G., Hradec, J. and Smits, P., *Modelling Inventory and Knowledge Management System of the European Commission (MIDAS)*, EUR 29729 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-02852-9, doi:10.2760/900056, JRC116242.

TOOL #62. MULTI-CRITERIA DECISION ANALYSIS

1. MAIN FEATURES

After having assessed their likely economic, social and environmental impacts, as well as distribution across stakeholders, **the impact assessment analysis compares the different options** with regard to their effectiveness, efficiency and coherence, as well as their compliance with the proportionality principle. As a consequence, here there is a need for a systematic, transparent and rigorous **integrative framework**.

In welfare economics and decision theory literature, the most traditional and widespread approaches for comparing options are **cost-benefit analysis (CBA)** and **multi-criteria decision analysis (MCDA)**. **Cost-benefit analysis is characterised by the attempt of measuring all effects of a policy option in monetary units.** CBA is grounded on market mechanisms; this implies that the analysis is based on the behaviour of individuals as consumers on the market. **Its main aim is thus to assess efficiency**[880]. Obviously, issues connected with actions outside of markets and behaviour of people different from the class of consumers cannot be considered[881]. To use only one measurement unit for incorporating a plurality of dimensions, objectives and values, implies a high risk of reductionism**. Multi-criteria decision analysis allows considering a wide range of assessment criteria, all of them shown in their original units of measurement;** there is no need to transform them in monetary terms as required by CBA. **MCDA** is particularly useful in case of complex interventions with diverse quantified impacts measured in different units and/or qualitative. This high degree of comprehensiveness and transparency is the most important success factor of multi-criteria methods.

MCDA is thus more comprehensive than CBA. CBA and MCDA can be considered as competitive methods *only* if **all** consequences of a policy decision can be correctly transformed into monetary values and efficiency is the only relevant policy objective. When a plurality of policy objectives exists, CBA can be used as one component of a MCDA framework dealing with the objective of efficiency in a consistent way.

Social multi-criteria evaluation (SMCE), which has been explicitly designed for public policy, also allows the capture of distributional consequences (e.g. in terms of stakeholder types, EU regions/countries or time) and trade-offs between dimensions (such as between some economic, social or environmental impacts, or between some families of criteria).

The **multi-criteria impact matrix** presents in a structured way the information on the various 'criterion scores', i.e. each element of the matrix represents the performance of a certain policy option according to each criterion (see Table 1). The various criterion scores can assess impacts by using both **quantitative** (e.g. as result of simulation models) and

[880] Alternatively, the extent to which objectives can be achieved for a given cost (cost effectiveness).

[881] "… there is such a long tradition in parts of economics and political philosophy of treating one allegedly homogeneous feature (such as income or utility) as the sole 'good thing' that could be effortlessly maximized (the more the merrier), that there is some nervousness in facing a problem of valuation involving heterogeneous objects, …And yet any serious problem of social judgement can hardly escape accommodating pluralities of values, …We cannot reduce all the things we have reason to value into one homogeneous magnitude." (A. Sen, The Idea of justice, The Belknap Press of Harvard University Press, Cambridge, Massachusetts, 2009, p. 239).

qualitative (e.g. results of participatory techniques) **information**. Some examples of criteria that can be all present simultaneously in a multi-criteria impact matrix include:

- "implementation cost for a Member State", "EU added value" (measured in euro), "administrative burden", "complementarity with other EU instruments", "avoidance of social exclusion" (measured in qualitative terms, e.g. very good, very bad or ++, --, …),
- "impact on labour force", "additional leukaemia risk" (measured by using probability distributions),
- "emissions reduction commitments" (measured in avoided CO2 emissions),
- "share of renewable energies" (measured in megawatt-hours of installed capacity),

		Options			
Criteria	**Units**	$\mathbf{a_1}$	$\mathbf{a_2}$	$\mathbf{a_3}$	$\mathbf{a_4}$
$\mathbf{g_1}$		$g_1(a_1)$	$g_1(a_2)$	.	$g_1(a_4)$
$\mathbf{g_2}$		.	.	.	.
$\mathbf{g_3}$		.	.	.	.
$\mathbf{g_4}$		.	.	.	.
$\mathbf{g_5}$		.	.	.	.
$\mathbf{g_6}$		$g_6(a_1)$	$g_6(a_2)$	.	$g_6(a_4)$

Table 1. Example of a multi-criteria impact matrix

The current practice for comparing policy options in Commission IAs very often builds such a matrix [882], thus the **use of standard definitions** from decision theory literature is recommendable (see Figure 1).

- **Dimension** is the highest hierarchical level of analysis. For example, in IA studies, this refers to the general categories of economic, social and environmental impacts.
- **Objective** indicates the desired direction of a policy change. For example, within the economic dimension, economic growth has to be incentivised; within the environmental dimension, the contribution to the EU's climate change commitments in the context of COP 21 has to be maximised; in the social dimension, the fairness should be maximised.
- **Criterion** is an empirical indicator that associates each policy option with a variable indicating its desirability according to its expected real-world consequences. Any objective may imply a number of different criteria. A classic example in the economic dimension might be GDP, saving rate and inflation rate inside the objective "growth maximisation".
- **Criterion score** is an assessment of the impact according to a given criterion with reference to each single policy option. Criterion scores can be both qualitative and quantitative; uncertainty can also be included. These assessments are the ones presented in the elements of the impact matrix (as in Table 1 above).

[882] Examples in published IA studies can be found in Munda (2017).

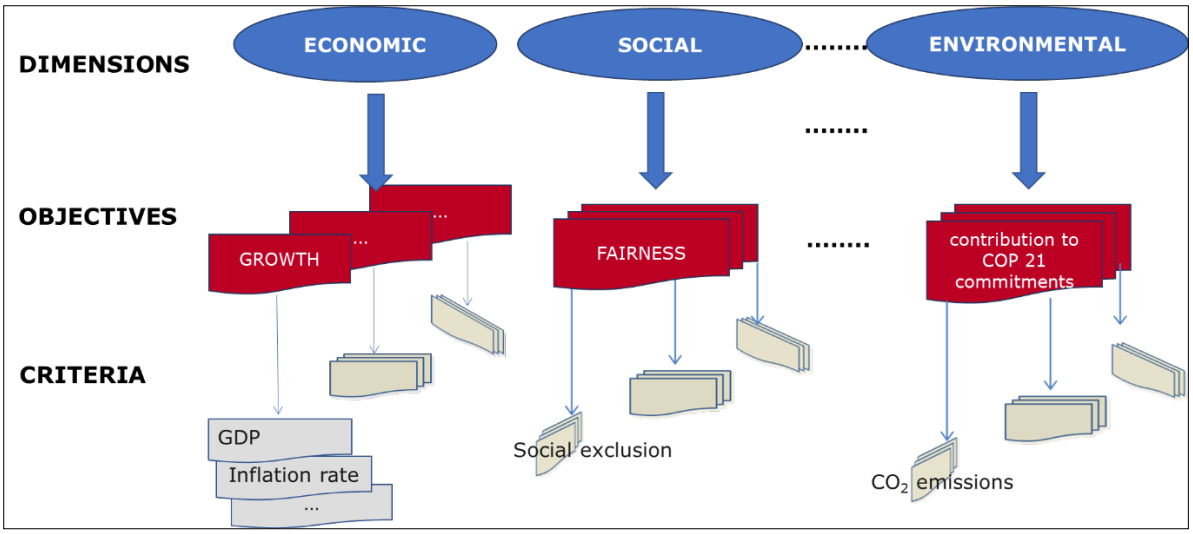


Figure 1: Example of the hierarchical structure for the application of SMCE to each option.

2. PRACTICAL IMPLEMENTATION STEPS

The application of SMCE is **not particularly time consuming**, since it **formalises in a consistent and efficient way a process that often is already done in the current practice of IA**. Many IA studies present the results in a form of an impact matrix.

Box 1. The main steps to implement a Social Multi-Criteria Evaluation framework

1. Selection of the relevant social actors for the problem at hand.
2. Definition of **social actors' values, desires, and preferences**. Stakeholder consultation methods described in Chapter 7 can be used to deal with both steps 1 and 2.
3. Generation of **evaluation criteria** to represent social actors' needs, preferences, and desires.
4. Construction of **the multi-criteria impact matrix** synthesising the performance of each option according to each criterion. In practice, this implies the integration in a coherent and integrative framework the results of the various sectorial models computing the various impacts (e.g., on economy, environment, health, energy, etc.).
5. Construction of an **equity impact matrix**, identifying all the distributional consequences of each single option in terms of stakeholder types, EU regions/countries or time (e.g. by considering consequences on future generations explicitly).
6. Application of a **mathematical procedure** to rank all the policy options in a consistent way. Multi-criteria mathematics solves the standard objection that the aggregation of apples and oranges is impossible in a definitive way. This makes the overall IA study much more defensible in comparison with the use of simple qualitative reasoning to isolate the most preferred option.
7. **Sensitivity and robustness analysis** checks if the ranking of policy options is stable and determines which of the input parameters influence more the output. Local sensitivity analysis looks at a) the exclusion/inclusion of different criteria and dimensions and b) changes of the weight of dimensions, criteria, or social actors; these are changed one per time. Global sensitivity analysis focuses on all the possible combinations of criterion

weights; all parameters are changed simultaneously (see also Tool #65 (*Uncertainty and sensitivity analysis*)).

In operational terms, two points deserve particular attention:

1) The **use of weights** can be a very sensitive issue. A reasonable practice can be to start by giving the same weight to each dimension and then splitting each weight among the objectives and criteria of any dimension proportionally. Of course, one could assume that some dimensions are more important than other ones, and thus their weight should be higher, but this should be justified based on strong and transparent ethical, scientific, institutional or legal arguments. Stakeholder processes can inform the weights attributed. **Sensitivity or robustness analyses** have to check the consequences for the final ranking of these arguments; they are thus a way to improve transparency on the assumptions introduced in an IA study.

From an operational point of view, the support of a **software tool makes all required computations very quick**. For example **SOCRATES**, developed by the JRC Competence Centre on Modelling of the JRC, is explicitly designed for IA problems. SOCRATES helps in structuring IA problems in the hierarchy dimensions, objectives, criteria and makes the weighting relations transparent, by also allowing for sensitivity and robustness analysis. SOCRATES also allows to analyse all the distributional consequences of each option.

3. ADDITIONAL INFORMATION

- Arrow, K.J., Raynaud, H. (1986) Social choice and multicriterion decision making. M.I.T. Press, Cambridge.
- Figueira, J., Greco, S. and Ehrgott, M. (eds.) (2016) Multiple-criteria decision analysis. State of the art surveys. Springer International Series in Operations Research and Management Science, New York.
- Munda G. (2017) A social multi-criteria framework for ex-ante impact assessment, EUR 28752 EN, doi 10.2760/909528
- Munda, G. (2008), Social Multi-Criteria Evaluation for a Sustainable Economy, Springer, Heidelberg, 2008 – also available from the Commission libraries.

TOOL #63. COST-BENEFIT ANALYSIS

1. MAIN FEATURES

Social investment has positive consequences for economic growth and societal well-being in general. However, as in any category of economic decisions, social investments have to consider their opportunity cost and try to answer difficult questions such as: how much of the national budget has to be devoted to environmental protection or education? Is it better to invest in pre-schooling or in universities? Is it better to invest in renewable energy or in pollution control? Economists try to answer these questions by considering the costs and benefits linked to each option; **cost-benefit analysis (CBA) is characterised by the attempt of measuring all effects of a policy option in monetary units**; it focuses on the selection of the option that exhibits the highest net benefit, which is considered as the most efficient one.

CBA assumes that all impacts of a policy can be monetised, and efficiency is the only relevant policy objective. Instead, **Multi-criteria decision analysis (MCDA) allows considering a wide range of assessment criteria**, which can be included in their original units of measurement (see Tool #62 for the relation between CBA and MCDA).

The **time and resources** needed for the development of a cost-benefit analysis depends a lot on the procedures used to monetise costs and benefits. This choice should be proportionate, and it depends on the extent of the impacts and available data. If the analysis affects several markets, directly or indirectly, one could opt for a general equilibrium approach (see Tool #61 (*Simulation models*)), but this needs more preparation and specialist expertise. On the contrary, a CBA does not need specialist software for the sequence of straightforward manipulations. However, monetisation can be very complex.

In practice, most impact assessments compare the costs and benefits of different policy options, using a mixture of data: monetary units; or quantified in a non-monetary way (e.g. ton of emissions/year); or qualitative, non-quantifiable data. This reflects the difficulties of monetising or quantifying all the relevant economic, social, and environmental impacts. What is important is that policymakers consider as solid an evidence base as is possible in a proportionate way. As such, most impact assessments are a blend of CBA and MCDA.

2. PRACTICAL IMPLEMENTATION STEPS

The **application** of cost-benefit analysis requires the following **main steps**:

1) Identification and monetisation of costs and benefits

An overview of the typologies of costs and benefits is presented in Tool #56. In CBA, the main effort consists in trying to apply the right valuation techniques to transform everything into money terms. Different techniques exist, and it is not always clear which one is most appropriate for a certain real-world problem. The principal ones are: contingent valuation, the travel cost method, hedonic pricing, and the shadow project approach. Among these only **contingent valuation** is universally applicable. The aim of contingent valuation is to elicit valuations (or 'bids') which are close to those that would be revealed if an actual market existed. Respondents say what they would be willing to pay or willing to accept if a market existed for the good in question. To determine the value of intangible goods and services, economists try to identify how much people would be willing to pay (willingness to pay, or

WTP) for these goods in artificial markets. Alternatively, the respondents could be asked to express their willingness to accept (WTA) compensation.

2) *Selection of the relevant time horizon and social discount rate*

Most new policies or projects result in **costs and benefits** that arise **at different times**. Typically, while costs tend to be concentrated earlier in the implementation, benefits tend to occur later in time (for example, building a new railway line has an immediate cost but provides benefits for many years in the future). The social discount rate is used to compare costs and benefits that occur in different time periods from the point of view of society (see Tool #64 (*Discount factors*)). **CBA results may vary a lot according to the discount rate chosen** (see Box 1). High discount factors tend to give a low value to future costs and benefits, thus prioritising close benefits and shifting costs to future generations. Here a sensitivity analysis is very useful tool to guarantee transparency.

3) *Choice of a mathematical aggregation rule*

In *empirical applications,* the most common mathematical aggregation rules are **net present value (NPV)** and the **benefit-cost ratio (BCR)**. The NPV is:

$$NPV = \sum_{i=0}^{i=n} \frac{B_i}{(1+r)^i} - \sum_{i=0}^{i=n} \frac{C_i}{(1+r)^i}$$

Where C_i and B_i are respectively the Costs and Benefits in a given year i over the policy/project lifetime of n years (starting in year 0), and r is the discount factor. For economic acceptability, **NPV** must be positive; the higher the NPV the more desirable the option.

In the case of **BCR**, the ratio between benefits and costs is used, instead of their difference. In this case, for economic acceptability BCR must be higher than one; the higher the BCR the more desirable the option.

To **choose only one option** among a set of competitive alternatives, the decision can be made on the grounds of the highest NPV. Instead, for a **complete ranking**, the choice of the right aggregation method to use is more complex, since NPV and BCR provide different results to the ranking problem. In fact, while NPV is an indicator of the attractiveness of an option in absolute terms (thus the larger the difference between benefits and costs, the better), in BCR the attractiveness is independent from the scale of options considered. Finally, one should also note that both NPV and BCR are very sensitive to the discount rate chosen, which could also change the ranking of the policy options (see Box 1 for an illustrative example).

4) *Presentation of the impacts and formulating the judgement on the performance of existing public intervention or the comparison of the policy options*

The different types of costs and benefits which have been monetised should be presented together with qualitative information on non-monetised costs and benefits. They should be compared in terms of the various cost/benefit categories, net benefits and net present value, as well as distributional impacts on stakeholders.

5) *Checking the robustness of the results*

It is important to highlight that **no '*objective value free*' approach exists**; results provided by CBA depend on a number of assumptions. Full **transparency on the assumptions** introduced in any specific application should always be assured. In the case of CBA, it should be noted that all monetary valuation attempts suffer technical uncertainties such as: Which monetary valuation technique has to be used? Which time horizon has to be considered? Which social discount rate? Which mathematical aggregation rule?

Box 1. Example: effect of the choice of the discount rate on the ranking between policy options

Time	Option A	Option B	Option C
1	-100	-100	-100
2	10	30	50
3	20	30	60
4	30	40	50
5	40	40	50
6	50	40	20
7	60	40	20
8	70	40	20
Discount rate	**NPV**		
1%	**165 (1)**	148 (3)	160 (2)
5%	114 (2)	107 (3)	**125 (1)**
10%	69 (3)	70 (2)	**92 (1)**

Let us consider three options A, B, and C. They present costs only at time 1 and then cash flows are all positive. Option A present high positive inflows at more distant time periods than option B and C. As one can see, this characteristic creates a rank reversal of A with B and C when the discount rate becomes higher (in the table above, the ranks are indicated in parenthesis).

6) *Accounting for distributional and cumulative impacts*

Distributional (e.g. Member States, richer and poorer groups, SMEs) and cumulative impacts (e.g. future generation) should also be considered in a proportionate way. It is crucial be aware of disproportionate impacts in the cost benefit analysis and consider segmenting the population in smaller groups to identify and illustrate distributional impacts where relevant.

3. ADDITIONAL INFORMATION

- Study prepared by the Centre of European Policy Studies on assessing the costs and benefits of regulation, available at http://ec.europa.eu/smart-regulation/impact/commission_guidelines/docs/131210_cba_study_sg_final.pdf
- Pearce D.W., Nash C.A. (1989) The social appraisal of projects, MacMillan, London.

- Newnan, D.G., Eschenbach, T.G., Lavelle, J.P. (2019) - *Engineering Economic Analysis*, 14th Ed., Oxford University Press, Oxford.
- Luenberger D. G. (2012) – Investment science, Oxford University Press, Oxford

TOOL #64. DISCOUNT FACTORS

1. INTRODUCTION

Individuals have time preferences, i.e., the availability of EUR 1 now is worth more than EUR 1 in the future. This can be explained as money can be invested today to generate a higher value tomorrow, there is inflation, or one will be dead in some future.

To capture this phenomenon one can use a discount factor. If an individual invests EUR 1 now to have some revenues in 5 years and wants to be compensated for the fact that this individual is not buying any good or service today with this EUR 1. This compensation is measured by a discount factor. A **discount factor** is a numerical factor used to **convert monetary values in the future to 'present values'**, so that money flows can be compared over time. It measures the present value of one euro received in year t. It relates to the complementary concepts of interest rate, rate of return, opportunity cost of a project, or cost of capital. In practice, discounting is using an appropriate interest rate back-to-front.

Most new policies or projects result in **monetary flows or/and net social effects (social benefits minus costs) that arise at different times.** Usually, costs have to be incurred in the present, so that benefits may be obtained in the future. The process of discounting is used to compare these monetary flows or net social effects at a given point in time, usually when the decisions about future private investments or public policies have to be made.

Discounting requires an important choice regarding the appropriate discount factor or discount rate as small differences in the discount factor can result in large changes of the net present value and thus can influence the evaluation of the proposal. The higher the discount factor used in an evaluation, the more difficult is the approval of the investment or the policy proposal (as the future benefits are smaller when converted into the present time).

One can distinguish between the **market discount factors** which usually are to be used when comparing private sector investment projects and the **social discount factors** which usually are used when evaluating future cost and benefits with societal value that are related to public policies. If the flows are to be compared from the point of view of the society, one should apply the **social discount factor**. If discounting is done from the perspective of an individual project, or the private sector, the discount factor applied will be different (usually higher) than a social discount factor as capital business projects are typically riskier than the projects related to public projects[883].

The choice of an appropriate social discount factor for public policy is highly debated in the economic literature. The most important theories are **the social rate of time preference rate (SRTP)**[884] and the **social opportunity cost of capital (SOC)**[885]. Each approach has its advantages and disadvantages.

[883] Opportunity cost of capital is the rate of return that capital could be expected to earn in the best alternative investment of equivalent risk. It is usually calculated by examining the returns of different projects available measured in net present value.

[884] The social rate of time preference consists of three components (SRTP=$p+e*g$), where p is the pure time preference, e is the elasticity of the marginal utility of consumption, i.e. the percentage change in individuals' marginal utility corresponding to each percentage change in consumption; g is the expected growth rate of per capita consumption. The positive pure time preference is about preferring the present over the future (as discussed before).

2. SOCIAL DISCOUNT FACTOR FOR EU POLICIES

As EU policies and projects funded by the EU budget are public policies with societal value, the social discount factor[886] can be used to assess them. Both investment projects funded by the EU budget as well as evaluations of EU policies call for a longer-term rate to be used, as such policies or projects typically need time to roll out and become effective. A social discount factor should be as transparent, factual, and predictable as possible. For practical reason this tool proposes a single central rate for all projects.

The social discount factor recommended for EU policy analysis is 3% in real terms[887]. This value follows the recommendation of the DG REGIO Economic Appraisal Vademecum[888,889] for EU-funded projects in the period 2021-2027.

However, for discounting health impacts or environmental projects, it is a common practice in many countries to choose a lower rate[890]. A similar argument is made to choose a lower rate for long-term impacts or projects. Section 6 stresses the need for **sensitivity** and to use alternative lower and higher discount factors than the central 3% value, in order to assess the robustness of the analysis.

If you are dealing with costs and benefits expressed in nominal prices one needs to adjust for the inflation: e.g. with an inflation of 2% per annum, a 5% **nominal social discount factor** (3% real rate plus 2% to account for inflation) would be used.

A certain stability over time is recommended. That should be balanced against possible changes due to evolving economic and social conditions. Societal projects are intended to deliver (direct and indirect) positive externalities. Therefore, social discount factors could be different than strictly financial discount factors.

3. FINANCIAL DISCOUNT FACTOR

The financial discount rate is used to determine the present value of future cash flows in a discounted cash flow analysis. This helps determine if the future cash flows from a project or investment will be worth more than the capital outlay needed to finance the public project or investment in the present. The cost of capital is the minimum rate needed to justify the cost of a new project or investment; and the discount factor is the number that needs to meet or exceed the cost of capital. This financial discount factor merely looks to the financial transactions directly related to the project or investment. It does not consider whether the

[885] The social opportunity cost of capital (SOC) argues that resources are scarce in any economy. This implies that the public and private sectors compete for the same funds. As a consequence, public investments should guarantee at least the same net social benefit as the private ones; otherwise an improvement can be obtained by reallocating resources to the private sector.

[886] A number of methods can determine the investment rates and the social discount factor for EU policy analysis such as Market-based determination, Judgemental determination, Model-based estimation, and Sharpe-Lintner CAPM.

[887] Historically governments across the world have been using social discount rates ranging from 2% up to 15%, depending on the long-term average risk-free rate. Rates tend to have declined over time. Therefore, the recommended social discount factor for EU policies is lower than the 4% in the 2017 version of the 'better regulation' toolbox.

[888] https://ec.europa.eu/regional_policy/en/information/publications/guides/2021/economic-appraisal-vademecum-2021-2027-general-principles-and-sector-applications

[889] Catalano, G. *et al*. (2021, forthcoming). The social cost of capital: Recent estimates for selected countries. CSIL working paper

[890] E.g. DE, DK, FR, NL, NO, and UK uses 1.5% for health and.

project or investment generate any benefits elsewhere in the economy (e.g. environment, or increased productivity in economic sectors).

The financial discount factor recommended for EU analysis is 0.5%[891] in real terms. As the financial discount factor is used for financial cash flows and based on long-term bonds, the value will change in function of the financial markets.

4. COSTS FROM THE PERSPECTIVE OF PRIVATE CAPITAL

There is widespread consensus that the social discount factor is usually lower than the discount factor that should be used for individual companies or households, who are unable to diversify risk as effectively as the society as a whole. The social discount factor is only used, therefore, when looking at issues from the societal point of view. For example, a higher discount factor should be used when trying to assess the behaviour of a company in respect of an investment decision. This would essentially be the internal rate of return required to trigger an investment. For a business, a good proxy is the **Capital Asset Pricing Method**, which takes account of both the costs of capital and the riskiness of the investment.

In some cases, the **Weighted Average Cost of Capital (WACC)** could also be used[892] when it is important to ensure that the future benefits are sufficient to compensate the required investment as well as the costs of funding. This approach may also be used for instance to evaluate projects that are only partially funded by the EU budget and the public resources are used to attract private investments.

Higher discount factors may also apply for households when deciding on whether to make an investment due to a range of factors: such as finance costs and other behavioural constraints like split incentives (e.g. landlord/tenant), short time horizons, risk averseness, information asymmetries or other obstacles or barriers. Similarly, there are proposals to use a higher discount factor for poorer actors as they have a higher necessity to satisfy today's needs compared to richer peers[893].

The discount rate used when deciding whether to invest may be different to the actual cost of financing as it includes other factors, barriers, or risks. E.g. for a firm, the cost of financing would be the weighted average cost of capital. However, hassle or transaction costs are a valid cost category and can be added to the total discount factor.

In a single analysis, different discount factors can be used. The policy measure itself can from a social point of view use a lower social discount factor; private actors may face higher discount factors; and a distinction may be made for income groups. However, while conceptually different discount factors can be justified, it is not straightforward to determine by how much these discount factor can differ. Therefore, all discount factors used should be transparent, and, if relevant, alternative discount factor values need to be explored to increase the robustness of conclusions.

---

[891] The figure of 0.5% is calculated in June 2021, using a 3 year moving average on the Euro area government bonds with 10-year maturity.

[892] The Weighted Average Cost of Capital consists broadly of a risk-free rate plus the Beta for the sector times the equity risk premium. Its value is not affected by a firm's choice between chosen equity and debt funding to fund investment.

[893] https://www.oxera.com/insights/agenda/articles/social-discount-rates-inequality-and-the-long-term/

5. THE MECHANICS OF DISCOUNTING: NET PRESENT VALUES (NPV)

Calculating the present value of the *difference* between the benefits *and* the costs provides the **Net Present Value (NPV)** of a policy option.

Box 1. Formula for the determination of Net Present Value

$$NPV = \sum_{i=0}^{i=n} \frac{B_i}{(1+r)^i} - \sum_{i=0}^{i=n} \frac{C_i}{(1+r)^i}$$

Where the Costs and Benefits in a given year i are C_i and B_i respectively over the policy/project lifetime of n years (starting in year 0), while r is the chosen discounting rate.

Box 2. Determination of present values using a social discount factor of 2.5% (and 4% for comparison)

The factor to discount present values is given by the formula below, where r is the discount factor and n is a future year:

$$Discount\ factor\ in\ a\ future\ year\ n = \frac{1}{(1+r)^n}$$

As an example, the present value of €1000 to be obtained in the future 5 years is shown below:

	Year					
	0	1	2	3	4	5
Present Value with 2.5%	€1000	€976	€952	€929	€906	€884
Present Value with 4%	€1000	€962	€925	€889	€855	€822

Box 3. The calculation of NPV for two competing policy options with 3%

Alternative projects A and B are both expected to improve the functioning of an organisation.

Option A: requires €10 million in capital costs initially in order to realise benefits of €2.5 million per annum in the following 4 years.

Option B: requires €5 million in capital costs initially to realise benefits of €1.5 million per annum in the following 4 years.

	Year	0	1	2	3	4	**NPV**
	Discount factor	1.0000	0.9709	0.9426	0.9151	0.8885	
Option A							
	Costs (€ m)	10	0	0	0	0	
	Benefits (€ m)	0	2.5	2.5	2.5	2.5	
	Benefits less costs (€ m)	-10	2.5	2.5	2.5	2.5	
	Present value (€ m)	-10.00	2.43	2.36	2.29	2.22	**-0.71**

Option B							
	Costs (€ m)	5	0	0	0	0	
	Benefits (€ m)	0	1.5	1.5	1.5	1.5	
	Benefits less costs (€ m)	-5	1.5	1.5	1.5	1.5	
	Present value (€ m)	-5.00	1.46	1.41	1.37	1.33	**0.58**

Project B realises a positive NPV of €0.58 million whereas Option A has a negative NPV of - €0.71 million. Therefore, Project B is preferable.

The net present value can be used to distinguish between two competing policy options. If the NPV is negative, this means that the future benefits are not sufficient to compensate the costs incurred given the selected discount factor (or required rate of return). The project with the higher NPV is to be selected. Boxes 3 and 4 compare the calculation of the NPV for 2.5% and 4%, respectively. While the order of the options does not change in these two examples, it can be observed that a lower discount rate leads to a higher NPV, and, hence, favours long-term investments.

Box 4. The calculation of NPV for two competing policy options with 4%

Alternative projects A and B are both expected to improve the functioning of an organisation.

Option A: requires €10 million in capital costs initially in order to realise benefits of €2.5 million per annum in the following 4 years.

Option B: requires €5 million in capital costs initially to realise benefits of €1.5 million per annum in the following 4 years.

	Year	0	1	2	3	4	**NPV**
	Discount factor	1.0000	0.9615	0.9246	0.8890	0.8548	
Option A							
	Costs (€ m)	10	0	0	0	0	
	Benefits (€ m)	0	2.5	2.5	2.5	2.5	
	Benefits less costs (€ m)	-10	2.5	2.5	2.5	2.5	
	Present value (€ m)	-10.00	2.40	2.31	2.22	2.14	**-0.93**
Option B							
	Costs (€ m)	5	0	0	0	0	
	Benefits (€ m)	0	1.5	1.5	1.5	1.5	
	Benefits less costs (€ m)	-5	1.5	1.5	1.5	1.5	
	Present value (€ m)	-5.00	1.44	1.39	1.33	1.28	**0.44**

Project B realises a positive NPV of €0.44 million whereas Option A has a negative NPV of - €0.93 million. Therefore, Project B is preferable.

6. SENSITIVITY ANALYSIS

To use one constant social discount factor across impact assessments and evaluation is important to ensure **coherence and comparability**; however, one should remember that

discounting at even modest rates reduces the value of monetary flows almost to zero over long time periods. This can be criticised because it excludes future generations from consideration in today's decisions. For example, let us consider a policy option with a non-discounted difference between benefits and costs of EUR 1 000 000. The higher the discount factor is, the quicker the net present value (NPV) becomes very low; for example, at 10% it is close to zero in 50 years, while at 1% or 2%, it is still positive after 150 years.

Discounting is a concept developed with the objective of being able to compare different cash flows between different time periods. The selection of the discount factor should be clearly motivated. **Sensitivity analysis of the effects of the applied discount factor (whether social or private) is highly desirable for assuring transparency**. Therefore it is recommended to have alternative calculations with sufficiently higher and lower values (up to +/- 1% at least).

7. CORRECTING FOR INFLATION

To consistently compare prices or macroeconomic variables (e.g. GDP) over time it is necessary to adjust observed prices for inflation over time (or to deflate). The nominal value are the current prices that are observed at a time. The constant prices are in real value, i.e. they have been adjusted for inflation with a price index. A price index reflects the inflation over time in comparison with a reference year.

$$\frac{\text{Current Price}}{\text{Price Index (decimal form)}} = \text{Constant Price}$$

8. ADDITIONAL INFORMATION

- Frederick S., Loewenstein G., O'Donoghue T. (2002) - Time Discounting and Time Preference: A Critical Review, Journal of Economic Literature, Volume 40, Number 2, 1 June , pp. 351-401(51).
- Harrison, M. (2010). Valuing the Future: the social discount rate in cost-benefit analysis, Visiting Researcher Paper, Productivity Commission, Canberra.
- HM Treasury (2018). "The Green Book, Central Government Guidance On Appraisal and Evaluation".
- Mohanram and Gode (2013). Removing predictable analyst forecast errors to improve implied cost of equity estimates.
- OECD (2018), Cost-Benefit Analysis and the Environment: Further Developments and Policy Use, OECD Publishing, Paris.
- Stern, N. (2014). The Economics of Climate Change. The Stern Review, Cambridge University Press.

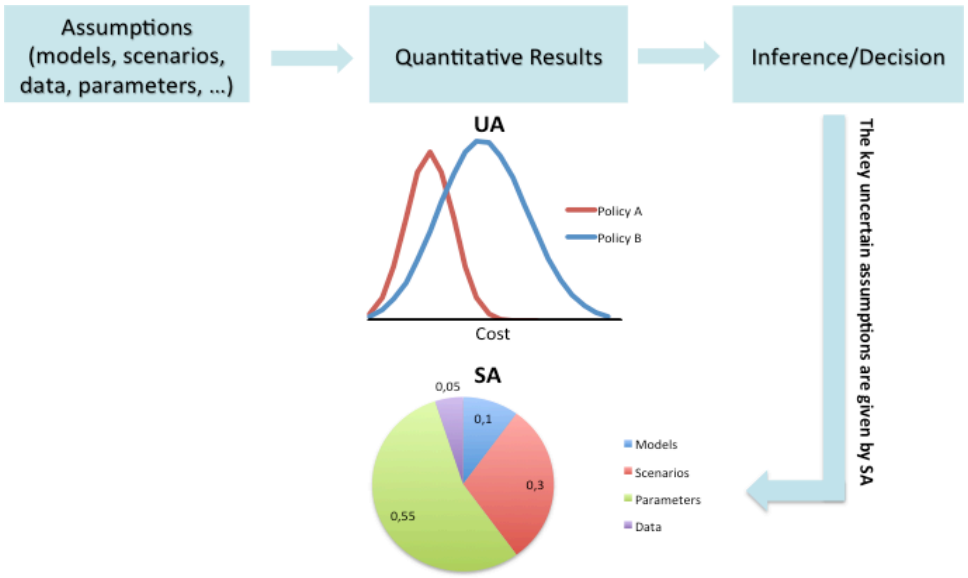
TOOL #65. UNCERTAINTY AND SENSITIVITY ANALYSIS

1. MAIN FEATURES

What	**Uncertainty analysis** aims at quantifying uncertainties in model results provided to the decision-makers due to uncertain assumptions/inputs**. Sensitivity analysis** allows identifying the uncertain assumptions mostly responsible for uncertainty in model results.
Why	A **transparent** and **high-quality impact assessment** should acknowledge and, to the extent relevant or possible, attempt to quantify the **uncertainty in results** as it could change the ranking and conclusions about the policy options.
How	**Assessing** the uncertainties in model results by propagating model input uncertainties through the model and **inferring** a posteriori the relevant uncertain inputs by subsequent statistical analysis.

Simulation models are extensively used for impact assessments (see Tool #61). When preparing and running these models, the current state of knowledge and all available data should be considered. Good and transparent practice in providing evidence for policy support requires that uncertainty be quantified and considered as much as possible (uncertainty analysis).

Box 1. The role of uncertainty analysis and sensitivity analysis in decision-making



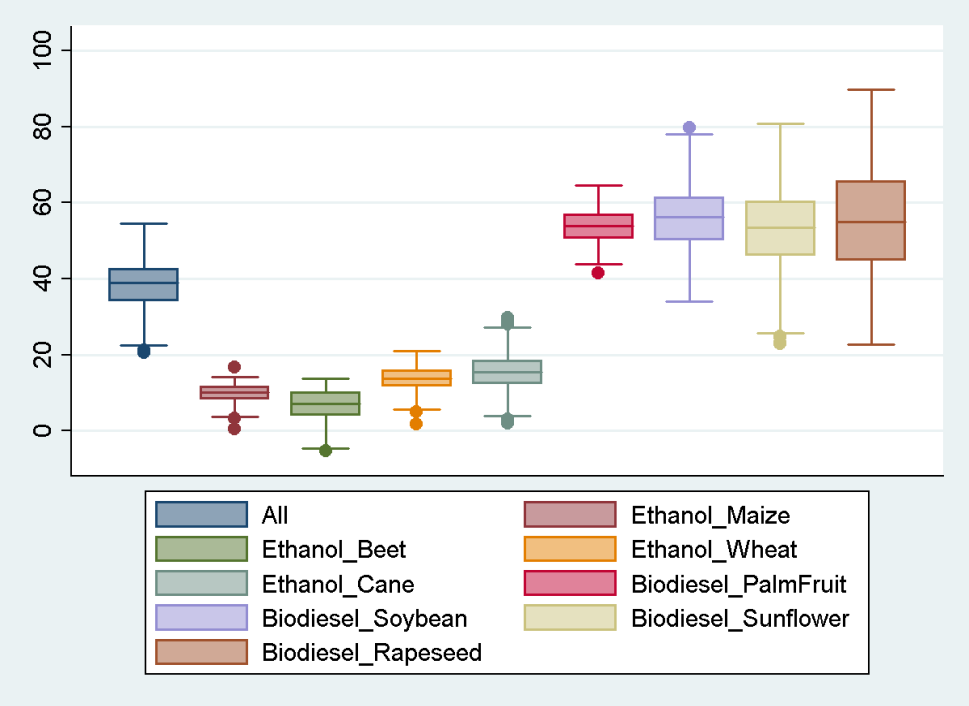
Different assumptions provide different results and can subsequently lead to different decisions. It is important to consider all plausible assumptions and eventually identify the key uncertain assumptions that drive the policy decision.

Uncertainty analysis aims at quantifying uncertainty in a model output using the propagation of uncertainty in the input variables. **Sensitivity analysis** is about understanding how the uncertainty in model results can be attributed to the different sources of uncertainty in the model inputs.

Uncertainty analysis can either i) point to a clear positive impact of a policy option compared with the others, or ii) fail at providing a clear decision because the uncertainties are too large. The first case leads to a robust decision because it means that, despite the uncertainties, the impact of the preferred policy option is consistently the better choice. In both cases, performing sensitivity analysis will allow to highlight which key uncertain variables (related to data and/or knowledge) are responsible for the uncertainty in policies' impact.

Box 2. Monte Carlo analysis to illustrate the range of uncertainty of ILUC GHG factors (IA on biofuels and indirect land use change (SEC(2012) 343)).

The IFPRI-MIRAGE-BioF model was used to model the consumption of biofuels used in the EU and to estimate the emissions of greenhouse gases associated with indirect land-use change for a range of biofuel feedstocks.



Results of the Monte-Carlo analysis: estimated indirect land-use change emissions (gCO_2/MJ)-under scenario of current trade policy. The bars indicate 1st and 99th percentile, while the boxes are 25th and 75th percentiles.

On the one hand, sensitivity analysis identifies which variables must be known more precisely to reduce the uncertainty in the results. Analysts can then try to refine their knowledge about the key variables (by expert judgments, data collection, measurement errors, etc.). On the other hand, sensitivity analysis informs the decision-makers about the key assumptions that drive the results and subsequently their decision (see Box 1).

Uncertainty analysis and sensitivity analysis are very informative ingredients of modelling for impact assessment and reporting the results in terms of range of values rather than a single value enriches the analysis and impact assessment (Box 2). Otherwise, one would give a spurious impression of accuracy. Furthermore, model results could also be accompanied with charts or indicators representing the importance of the key variables for each quantitative outcome.

It is important to stress that uncertainty analysis and sensitivity analysis are intended to be **problem-oriented** and not model-oriented. This means that uncertainty and sensitivity analyses may provide different results for different impact assessments even though the same models are employed. This is because the assumptions and data might differ from one study to another.

Undertaking uncertainty analysis and sensitivity analysis is likely to require extra computational, human or financial resources to be deployed during the impact assessment. These resources may not be routinely available for particularly complex models. Nonetheless, those undertaking modelling studies should attempt to identify the key drivers of their results. The use of emulators (or meta-models) could reduce the computational burden associated with sensitivity analysis.

There are different ways to quantify uncertainty. 'Local' uncertainty analysis and sensitivity analysis examine the variation in the model output by changing one input variable at a time, usually to the minimum and maximum plausible values. This **'one-at-a-time' (OAT) approach** is commonly used as it requires less resources, but it can be inaccurate or insufficient.

A 'global' approach allows for the simultaneous exploration of all sources of known uncertainty and can capture nonlinearities and interactions between model inputs. In global uncertainty and sensitivity analysis (GSA), probability distributions are assigned to uncertain model inputs. This uncertainty is then propagated through the model by running it repeatedly with different input values, which provides probability distributions of the model output of interest.

2. PRACTICAL IMPLEMENTATION STEPS

Uncertainty and sensitivity analyses imply a carefully planning during the model design and execution for impact assessment (see Tool #4 (*Evidence-informed policymaking*) and the introduction to Chapter 8). A timely evaluation of the cost and feasibility of these analyses is primordial.

Box 3. The basic steps to performing global uncertainty and sensitivity analysis

1. **Define the variables of interest for the analysis**. These variables should include the most relevant outputs for the impact assessment.
2. **Identify all model variables that are affected by uncertainty** in consultation with experts and stakeholders as appropriate. Inputs can be of various nature, i.e. scalar variables, time series or spatially distributed maps.
3. **Characterise the uncertainty for each selected input** by assigning a probability distribution using all available information such as experiments, estimations, physical bounds considerations and expert opinion. Extended peer-review should be considered. This crucial step may require significant resources.
4. **Generate a sample** from the previously defined probability distributions. The sample is a matrix which specifies the input values to use for each model run. The sample is generated so as to explore the full extent of uncertainty and is based on the input distributions specified in the previous step. Software packages are available for this.
5. The model is run many times using the sampled input variables for each model run as

identified in the previous step. **For each run, the value of the output variables** of interest **is recorded.**

6. The results of the model runs are then used to **estimate sensitivity indices, as well as uncertainty** in the model output.

Sensitivity analysis can be complicated, impractical or infeasible. E.g. large models require sufficient computing power and may take a long time to run. There may also be large numbers of uncertain model inputs, and correlations between input variables.

Box 4. A simple example of sensitivity analysis

A model is built to estimate the potential economic cost of a chemical accident at a proposed plant in a European region, including trans-boundary effects. It examines the number of people and businesses living within a certain radius and estimates the total value of lost property and life corresponding to different classes of explosion or fire.

Applying sensitivity analysis, the output variable of interest is the total cost of the damage. Uncertain inputs include medical costs per individual, total population within the impact radius, the size of the impact radius, and the assumed proportion of people and businesses affected, among others. Using expert opinion and available statistics, probability distributions are assigned to each variable, and a sample is constructed consisting of some thousand runs of the accident model. The sample is used to run the model, and the resulting output vector is used to estimate sensitivity.

It is found that, with 95% confidence, the estimated cost is within €2Bn to €20Bn. Furthermore, the most influential input variable is the stock of flammable material, causing 38% of the variance in the cost, followed by engineering variables accounting for 15% of the variance, with a set of meteorological parameters (wind speed and direction) accounting for most of the remaining variance.

It should be noted that sensitivity analysis addresses uncertainties that can be quantified (Box 3). In some cases, i.e. when a deeper assessment of the framing of the analysis is needed, or where there is a major disagreement among stakeholders about the nature of the problem, the analysis can be extended to **sensitivity auditing**.

Sensitivity auditing is a wider consideration of the effect of all types of uncertainty, including structural assumptions embedded in the model, and subjective decisions taken in the framing of the problem. In general, sensitivity auditing stresses the idea of clearly communicating the extent to which particular models can be used to support policy decisions and their results can be trusted, considering as much as possible all forms of potential uncertainty, and to anticipate criticism by third parties. In particular, one should avoid giving the impression of false confidence by 'quantification at all costs'. In some cases there is simply not enough data, or the process is too complex, to give a meaningful quantitative prediction.

3. ADDITIONAL INFORMATION

The JRC (Competence Centre on Modelling) has developed an online tool for uncertainty analysis and sensitivity analysis, available to Staff of the European Commission at https://web.jrc.ec.europa.eu/rapps/sensitivity/

- Azzini, I., Listorti, G., Mara, T. and Rosati, R., Uncertainty and Sensitivity Analysis for policy decision making, EUR 30432 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-24752-4, doi:10.2760/922129, JRC122132.
- Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana and S. Tarantola, *Global Sensitivity Analysis. The Primer*, Chichester: Probability and Statistics, John Wiley and Sons, 2008.
- Saltelli, A., Guimarães Pereira, Â., Van der Sluijs, J.P. and Funtowicz, S., 2013, What do I make of your latinorum? Sensitivity auditing of mathematical modelling, Int. J. Foresight and Innovation Policy, 9, 2/3/4, 213–234

TOOL #66. LIFE CYCLE ASSESSMENT

Sustainability dimensions (environmental, social and economic) should be considered in an integrated and holistic manner. By adopting life cycle thinking, impacts can be assessed:

- embracing all steps of value chains, namely of production and consumption systems (e.g. from extraction of raw materials to end of life/waste management);
- fostering comprehensiveness, e.g. entailing different kind of impacts;
- unveiling trade-offs and avoiding shift of burdens from one life cycle stage to another (e.g. from extraction to processing or processing to consumption phase); or across impact categories (e.g. improving on climate change while worsening in water use); or in terms of spatial and temporal resolution (e.g. shifting impacts from within the EU to other world regions or from current generations to future ones).

1. MAIN FEATURES

Life-cycle thinking is a broad concept that facilitates an integrated assessment of the benefits and the burdens in terms of environmental, social, and economic aspects, and can be applied at different scales, from products to regions and complex systems. The application of life-cycle thinking requires specific methods for impact quantification, such as Life Cycle Assessment (LCA) to assess environmental impacts, social LCA to evaluate social impacts, and Life Cycle Costing to assess direct and indirect economic impacts. The combination of these assessments methods provides a complete and comprehensive Life Cycle Sustainability Assessment. This tool focuses on LCA, i.e., on the environmental dimension; while many of the principles are analogous for the other sustainability dimensions. Similarly, this method can also be applied to the social dimension.

LCA is a holistic approach, which supports the integration of environmental sustainability into design, innovation and evaluation of goods and services (products). Being developed since the 1970, LCA is now a mature environmental assessment methodology, internationally standardised (ISO14040/44, 2006). LCA is mentioned as leading principle in many EU[894,895,896,897] and international[898,899] policies since the 90's[900]. Moreover, LCA is now central in several European Green Deal policy initiatives and beyond [901,902,903,904,905,906,907].

[894] Integrated Product Policy - Building on Environmental Life-Cycle Thinking. Communication from the Commission; COM(2003) 302

[895] Thematic Strategy on the Prevention and Recycling of Waste. Communication from the Commission; COM(2005) 666.

[896] A Sustainable Bioeconomy for Europe: Strengthening the connection between economy, society and the environment. COM(2018) 673 final.

[897] Building the Single Market for Green Products. Communication from the Commission; COM (2013) 196 final

[898] UNEP, 2004. Why take a life cycle approach? Paris, p 28

[899] UNEP-SETAC life cycle initiative; http://www.lifecycleinitiative.org/

[900] For an overview of the use of LCA in: Sala et al. (2021). The evolution of life cycle assessment in European policies over three decades. The International Journal of Life Cycle Assessment, 1-20.

[901] A new Circular Economy Action Plan For a cleaner and more competitive Europe; COM(2020) 98 final.

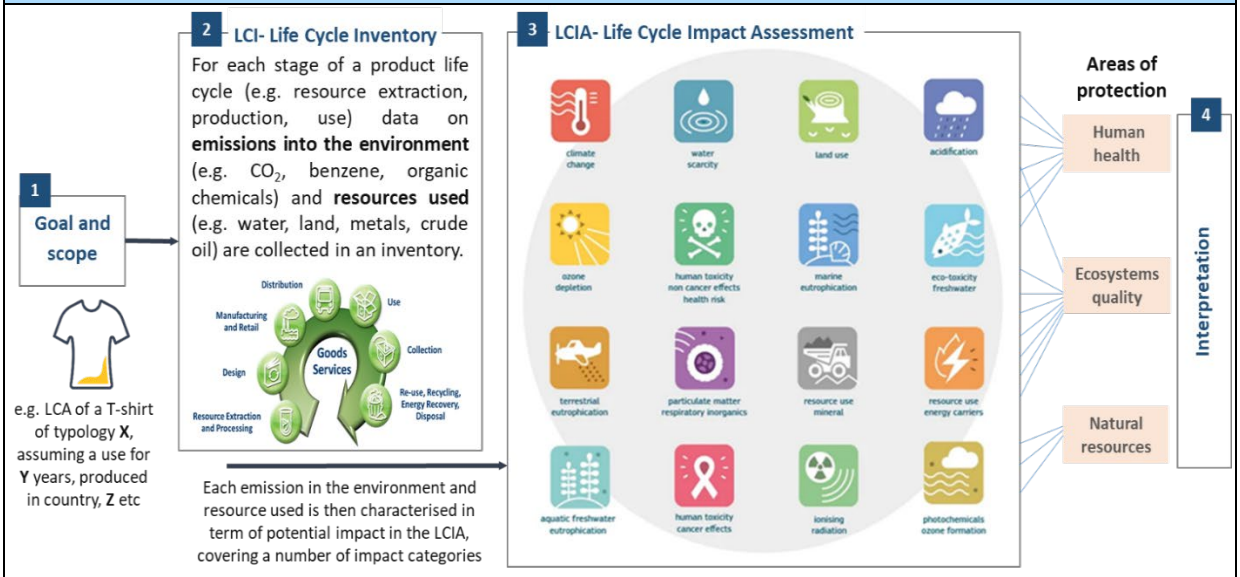
[902] A Farm to Fork Strategy for a fair, healthy and environmentally-friendly food system; COM(2020) 381 final.

LCA aims at an integrated environmental assessment of products (goods and services) and organisations along their supply chain, covering a wide variety of pressures and impacts associated with human health, ecosystem quality, and resources consumption. The method enables the comparison of the environmental impact of products by quantifying all inputs and outputs of energy and material flows in each stage along the entire life cycle. By applying a life-cycle approach, priorities and trade-offs can be identified more transparently thereby potentially resulting in more effective policies.

In an LCA, the natural resources (raw material, water, land) consumed and the emissions into air, water and soil are quantified in an inventory. The potential burdens associated to that inventory are then quantified using specific assessment models for each category of impacts, resulting in a number of impact indicators addressing different impact categories, e.g. climate change, eutrophication, ecotoxicity, land use-related impacts etc.

According to the ISO standard (ISO 14040), LCA consists of four phases (*Box 1*):

Box 1. Procedural steps of LCA



The basic scheme of a LCA study involves four phases: goal and scope, life cycle inventory (LCI), life cycle impact assessment (LCIA) and interpretation. The environmental impact categories within LCIA are those currently in use in the EU Environmental Footprint.

- **Goal and scope (1)** definition phase: definition of the aim(s) of the LCA study and description of the central assumptions and system characteristics (e.g. system boundaries, scenarios) in the assessment are described; the environmental impacts refer to a functional unit (e.g. the function provided by a car, a litre of milk) set as a reference quantity for the study, reflecting a specific product/service and its function, based on the defined goal and scope);

[903] Pathway to a Healthy Planet for All EU Action Plan: 'Towards Zero Pollution for Air, Water and Soil'. COM(2021) 400 final.
[904] Chemicals Strategy for Sustainability Towards a Toxic-Free Environment. COM(2020) 667 final.
[905] EU Biodiversity Strategy for 2030 Bringing nature back into our lives. COM(2020) 380 final.

Box 2. Examples of LCA results

The comparison of two products is presented by highlighting the relative performance in each Environmental Footprint category. For example, in the figure below the environmental impacts of two products A and B are shown, with the highest impact value set to 100% and the other option as relative to that. The analysis can be done also for assessing hotspots of impacts. The example highlights the relevance of each life cycle stage per each impact category for product A.

Impact categories	Product A	Product B	Product A life cycle				
			Raw material extraction	Manufacturing	Distribution	Use stage	End of Life
Climate change	100%	90%					
Ozone depletion	29%	100%					
Particulate matter	87%	100%					
Ionising radiation	70%	100%					
Photochemical ozone formation	25%	100%					
Acidification	91%	100%					
Eutrophication, terrestrial	86%	100%					
Eutrophication, marine	100%	96%					
Eutrophication, freshwater	100%	85%					
Ecotoxicity, freshwater	100%	92%					
Human toxicity, non-cancer	51%	100%					
Human toxicity, cancer	100%	88%					
Land use	100%	95%					
Water use	100%	89%					
Resource use, fossil	100%	87%					
Resource use, minerals and metals	100%	60%					

Very relevant impact
Relevant impact
Slightly relevant
Not relevant

- **Life Cycle Inventory (LCI) (2)**: data collection of resource use and emissions related to the chosen products/services for each process in each life cycle stage, from extraction of raw material to end of life;

- **Life Cycle Impact Assessment (LCIA) (3)**: resource use and emissions data collected in the LCI phase are translated into indicators that reflect impacts associated with human health, ecosystem quality, and resources consumption, covering different impact categories. This calculation is based on specific impact assessment models and characterisation factors, which represent the potential environmental impact per unit emission or resource consumption. The impact on different impact categories may then, be associated with three protection areas: human health, biodiversity (ecosystem quality), and natural resources. The EU Environmental Footprint method identified 16 impact categories and indicated the use specific models for assessing those impacts[908].

- **Interpretation (4)**: the outcome of the LCA calculation is interpreted in accordance with the aim defined in the goal and scope of the study. This step (illustrated in Box 2) is frequently performed in a decision-making context that requires political choices.

[908] EU recommended impact categories, models and factors are available at https://eplca.jrc.ec.europa.eu/EnvironmentalFootprint.html

LCA is a bottom-up methodology developed to capture in a systematic manner complex systems and supply chains. The more representative the data, the more robust is the LCA analysis. Methodological developments are aiming at improving over time LCA effectiveness in addressing cause-effect relationships in dynamic and rapidly evolving production and consumption systems. Some supply chains are well characterised (e.g., EU production), while others need further improvement (e.g., emerging products and global markets, like food and feed production in developing countries). EU and international efforts are ongoing to ensure that the large amount of data originating from different sources are reaching the best possible quality (in terms of accuracy, technological, temporal and geographic representativeness). Sensitivity and uncertainty analyses can be conducted to improve the robustness of the results.

Box 3. European Platform on Life Cycle Assessment (EPLCA)

The European Commission has acknowledged the potential of LCA as a decision-support method and has further refined LCA to a policy-decision context, considering the methodological and data-related challenges. The European Commission's JRC established the **European Platform on Life Cycle Assessment**, which represents the reference point for data and methods recommended at EU level to implement life cycle-based approaches. Complementary to the Platform, the **Life Cycle Data Network**, aims to provide an international basis for inter-operable, quality assured life cycle inventory data. It also provides a series of tools, guidelines, reference packages and format specifications to facilitate and harmonise the development of those data.

The Platform launched the **International Reference Life Cycle Data System (ILCD)** handbook. As LCA results could differ for a multitude of reasons including input data used, differences in approach, system boundaries, reference systems and different methods of allocation of the impacts between different products and co-products, this handbook provides a series of operational guidance documents for different types of LCA applications to support robust, replicable and transparent assessment.

The handbook constituted the basis for the development of the **Product Environmental Footprint (PEF)** and **Organisation Environmental Footprint (OEF)** methods. The two methods are annexes to the Commission Recommendation (2013/179/EU) on the use of common methods for measuring and communicating the life cycle environmental performance of products and organisations. These methods provide harmonised approach for multi-criteria environmental LCAs and build on international guidelines and ISO standards (e.g. ISO 14040, 14044, 14067, 14072). The two methods provide practical and prescriptive guidance for performing a more robust, consistent, reproducible and verifiable environmental assessment of products and organisations.

PEF and OEF can be used to substantiate green claims in line with the new Circular Economy Action Plan [909]. To enable comparisons within product groups and sectors, **Product Environmental Footprint Category Rules** (PEFCRs) and **Organisation Environmental Footprint Sector Rules** (OEFSRs) are developed, which provide more specific rules for performing the environmental assessment of product groups and economic sectors based on sector-specific organizations representing the EU and global markets. These rules (e.g. the choice of the functional unit) avoid methodological choices and increase he comparability and robustness of the results.

[909] A new Circular Economy Action Plan For a cleaner and more competitive Europe; COM(2020) 98 final.

2. PRACTICAL IMPLEMENTATION

LCA has been used in different applications related to policies, such as:

- **Support of EU policy development:** LCA is central in several EU policies, including the European Green Deal related initiatives[910]. LCA is mentioned across various policy domains, such as environmental product policy, renewable energy and waste policy. Some examples of LCA use in EU policy development are here reported:
    - Definition of emerging problems related to products and product supply chains, and new technologies, such as to evaluating the environmental performance of economic activities included in the EU Taxonomy[911].
    - Identification of policy options: e.g. the impact assessment of plastic bags directive has been based on a number of different LCA's.
    - Development of voluntary environmental product policy instruments, such as the definition of minimum environmental criteria (e.g. EU Ecolabel Regulation, Green Public Procurement).
    - Development of mandatory requirements for product (e.g. communication of carbon footprint for new batteries put into the market, according to the recent proposal for Regulation on Batteries[912]).
- **Enable comparisons within product groups and sectors**: In 2019, 19 Product Environmental Footprint Category Rules were finalised covering different products (such as batteries and accumulators, several food products – e.g. dairy, beer, packed water, household detergents, IT equipment) and 2 Organisation Environmental Footprint Sector Rules (copper production and retail).
- **Baselines of environmental impacts to test policy options**: The LCA-based Consumption Footprint and Consumer Footprint indicators [913] assess the environmental impacts of consumption at Member State and at individual (citizen) level respectively. These indicators can be used as baselines to assess the environmental benefits of policies (see e.g. the role for the monitoring of the circular economy[914]) and to monitor in a holistic manner the evolution of impacts in relation to consumption patterns. Moreover, impacts could be compared against planetary boundaries to assess the extent to which EU production and consumption are surpassing sustainability thresholds.
- **Supporting implementation of UN SDGs at product/service level** with LCA (e.g. in support to the achievement of the SDG 12 on responsible consumption and production; and assessing transboundary impacts of EU production and consumption on other world regions).

---

[910] Sala et al. (2021). The evolution of life cycle assessment in European policies over three decades. The International Journal of Life Cycle Assessment, 1-20.

[911] Regulation (EU) 2020/852 of the European Parliament and of the Council of 18 June 2020 on the establishment of a framework to facilitate sustainable investment, and amending Regulation (EU) 2019/2088. EU(2020) 852.

[912] Proposal for a Regulation of the European Parliament and of the Council concerning batteries and waste batteries, repealing Directive 2006/66/EC and amending Regulation (EU) No 2019/1020; COM (2020) 798 final.

[913] https://eplca.jrc.ec.europa.eu/sustainableConsumption.html

[914] A new Circular Economy Action Plan For a cleaner and more competitive Europe. COM(2020) 98 final.

3. ADDITIONAL INFORMATION FEATURES

- European Platform on Life Cycle Assessment https://eplca.jrc.ec.europa.eu/
- Sala et al. (2016). Life cycle assessment for the impact assessment of policies. Luxembourg. Publications Office of the European Union: ISBN 97892-79-64813-7
- Sala et al. (2019) Indicators and Assessment of the environmental impact of EU consumption. Consumption and Consumer Footprint for assessing and monitoring EU policies with Life Cycle Assessment, Luxembourg: Publications Office of the European Union, ISBN 978-92-79-99672-6
- Sala S. (2020) Life cycle assessment and evaluation of solutions towards sustainable development goals. In: Filho W. L. (Eds). Encyclopedia of the UN Sustainable Development Goals, Springer, ISSN: 2523-7403
- Sala, S., Amadei, A. M., Beylot, A., & Ardente, F. (2021). The evolution of life cycle assessment in European policies over three decades. The International Journal of Life Cycle Assessment, 1-20.

TOOL #67. DATA IDENTIFICATION FOR EVALUATION AND IMPACT ASSESSMENT

This tool gives an overview on the data collection required to produce evidence on policy impacts.

1. INTRODUCTION – THE NEED FOR EVIDENCE

Impact assessments aim to collect ex-ante evidence to assess the environmental, social, and economic impacts an intervention is expected to have. Evaluations aim to collect ex-post evidence to assess an intervention against the five evaluation criteria of (i) effectiveness, (ii) efficiency, (iii) coherence, (iv) relevance and (v) EU added value. The assessment is applied both to the intervention itself, as well as to related policies such as delegated and implementing acts. The ultimate aim of the impact assessment is to explore the best possible future action, while the ultimate aim of evaluations is to learn what worked from experience and to improve the policy. The present tool classifies data and asks relevant questions on 'what data to collect' for impact assessments and evaluations. Moreover, it provides guidance on how to design data collection at an early stage.

2. DATA ACROSS THE POLICY CYCLE

The following 11 questions recap some of the main data considerations, focusing on the obvious and less obvious links related to data among the three main simplified phases of the policy cycle, namely the ex-ante, the implementation and the ex-post phases. It is important to reflect on these questions throughout *all phases* of the policy cycle.

1. **What are the objectives and issues addressed by the policy**? What does the current intervention aim to change, i.e. what are its objectives? What are the main issues that the policy aims to tackle? What are their drivers and which ones is the policy going to intervene on?
2. **Why should it work?** All policy initiatives benefit from designing upfront the 'intervention logic', where for instance the rationale of the policy intervention, policy context, expected chain of effects on the intervention, and expected outcomes of the intervention (both intended ones and side-effects) are explicitly listed and discussed usually in an impact assessment. The use of such an intervention logic has the advantage to consistently guide actors through different phases of the policy cycle i.e. from design phase to implementation and evaluation with some learning/adaptation of the intervention logic allowed along the way.
3. **What will the success look like?** Given the intervention logic, one should choose appropriate indicators that would help measuring whether the intervention was successful or not. To the extent possible, the selected indicators should cover both intended effects and side-effects. This can feed into a plan on how to access data on these outcomes, the 'data plan'. The number of indicators of success should be limited to a manageable handful.
4. **Who is affected?** EU legislations affect individuals/firms/regions to a different extent. Answering this question can lead to a plan on how to obtain data on relevant individuals/firms/regions.

5. **What data plan is needed?** An early 'data plan' (see 3.3) is the biggest facilitator for any future evaluation. If done upfront, it efficiently guides a fit-for purpose ex-ante phase, informs the implementation phase and carries the blueprint for the evaluation phase.
6. **What type of impacts need to be considered?** Depending on the intervention to be evaluated all relevant impacts in addition to social and/or economic and/or environmental ones, need to be considered. In addition, for economic impacts, special consideration is required on compliance costs and in particular administrative costs for which the 'one in, one out' approach applies (see Tool #59 (*Cost estimates and the 'one in, one out' approach*)).
7. **What types of data need to be considered?** Stakeholder consultations provide data on *opinions*. Data from markets, balance-sheets of firms, registries (health, social security, unemployment) etc. are *observational data*. Observational data can be micro (having records for every single individual or firms, etc.) or macro (as in national accounts aggregating data over a specific group). Data on opinions complement observational data. These two types of data are complementary but not substitutes.
8. **What are the interactions, the trade-offs, the synergies?** What interactions are there between the policy initiative and other initiatives? Are these trade-offs, synergies or other interactions? What does this imply for the data plan, i.e. can one address data collection jointly with other parts of the package?
9. **Where should data come from? What is its quality**? Monitoring and evaluation clauses could mandate the (re)use of data that is already collected by local and national government (registry data, market data) or data from EU institutions. This way to access data is the most cost-effective. However, if new data is needed for evaluation purposes, monitoring and evaluation clauses should appropriately mandate their collection and sharing (see Chapter 5 on monitoring).
10. **Are ex-ante and ex-post impact indicators the same?** There needs to be consistency across phases of the policy cycle. Outcomes/impacts considered in the ex-ante and ex-post phases need to be measured by the same indicators. These outcomes should also be included in the monitoring phase and serve as a basis for the 'evaluate first' principle.
11. **Can one use monitoring data for evaluation?** There are synergies between monitoring and evaluation. Monitoring should aim to collect data on the impacts, which are usually identified in the ex-ante and gathered in the ex-post phases. If monitoring data is collected or accessed for individual/firms/regions with different treatment level or status this will allow having better monitoring information, where differences between groups will signal success of the intervention in a more robust way.

3. WHICH DATA TO COLLECT?

Not 'any data' can serve as evidence for an intervention. Data needs to be relevant in respect to: (a) the impact assessment and evaluation criteria (b) the specific intervention logic. Therefore, using 'the data that we have', however tempting it is in a situation where evidence is limited, should be assessed in view of its usefulness.

The best data for an impact assessments and evaluation is linked to the policy objectives and intervention logic. If choice exists, it is preferable to collect more granular data, as this facilitates the use of causal methods (see section 1).

3.1. From the intervention logic to measuring impacts

A simplified intervention logic for an evaluation is outlined in Tool #46 (*Designing the evaluation*), and the use of an intervention logic / problem tree for impact assessments is highlighted in Tool #11 (*Format of the impact assessment report*).

Based on the intervention logic, one should be able to identify data for:

i. *Policy inputs*; these are the means to achieve the EU policy objectives of a given intervention. In spending programmes, it is important to record how much financial resources were given to any single recipient of the funds. Often policy input indicators are collected by managing authorities or by the Commission. It is hence important to record and make accessible this type of data;

ii. *Outputs*; this is what is expected to be generated directly by the policy intervention;

iii. *Results/impacts*; these are what the given intervention expects to achieve in the medium to long term.

iv. *External/contextual factors;* these are factors that can influence the results/impacts of the intervention in addition to policy inputs; external/contextual factors are important. For example, to measure the impact of most social policies one needs to account for the macroeconomic conditions related to the business cycle (external/contextual factors).

v. *Other overlapping policies*; the coherence with other policies can affect the comparison point and degree of analysis possible.

In order to frame the question 'what data to collect?' one needs to consider the intervention's expected results/impacts, since they dictate the next steps on data identification and collection.

A complementary way to approach the question 'what data needs to be collected' is to focus on the main actors that the policy is targeting. For instance, if a policy initiative is trying to guide citizens to favour sustainable products by regulating labelling of goods, one could consider as relevant indicators the level of sales of sustainable goods on the side of consumers, and possibly the level of costs incurred by firms to label the goods. This may imply the need to collect data on sales and on firms' costs, probably from different sources.

One consequence of this observation is that, while *all* EU policies aim at improving the welfare of citizens, evaluations need to stay *close to the specific aims* of the given policy, starting from providing evidence on behaviour of the main actors that the policy aimed to influence.

3.2. Unintended effects

When answering the question 'what data to collect' unintended or side-effects of a policy intervention need also to be considered. Using the labelling example above, the requirement for new labelling of goods may induce higher production costs, which may (at the extreme) force some companies to reduce their manpower. This is an example of unintended effects or side-effects of a policy intervention.

3.3. Data availability and data plans

If some of the needed data is not already available, what can be done? A lot can be and should be planned in advance, for instance using a carefully constructed stakeholder consultation strategy. Early action is often the cure to many of the data limitations encountered later in the EU policy cycle.

The first principle is to start planning a clear monitoring structure from the impact assessment / design phase (see Tool #43 (*Monitoring arrangements and indicators*)). If it is recognised that the right data is not currently available, then the design of the legal text may envisage ways how to make it available next time an evaluation is to be carried out. One way to do this is to insert specific monitoring and evaluation provisions. In addition, when there is a lack of data make sure to explore all of the Commission resources available. That includes, but is not limited to, checking internally for data from previous evaluation and studies, data from the work of other DGs such as Eurostat or EU decentralised agencies and other EU bodies, as well as engaging with data platforms such as RegHub, Fit4Future or EESC in the quest for data.

Is some data better than no data? Yes, but not any data. If the available data at hand is related to the evaluated policy in the sense of being linked to the intervention logic, then it should be used. However, even large amounts of data that are only remotely linked or not informative about the given policy are useless. So it is important to think ahead and plan collecting data from the beginning of the policy cycle.

3.4. Data triangulation

If the data plan is well done, one may find different sources of data that try to measure the same phenomenon. Data triangulation is a way to validate the quality of different sources of data. For instance, one may cross-check the opinions collected in consultations with statistical data on the same issue. Should they not match, further investigation is required on the possible reasons.

3.5. Impacts ex ante and ex post

The outcome variables (impacts) of an evaluation should be the same as the ones considered in the impact assessment, and the outcome variables in an impact assessment should in the best-case scenario build on a previous evaluation, when available. The different evidence base gathering processes of the policy cycle should help improve each other and is another reason for a careful and thoughtful design of the monitoring of an intervention.

If something cannot be measured in the evaluation, maybe it should not have been included as impact in the impact assessment or the relevant monitor indicator was missing. This consideration may limit the number of impacts considered in the subsequent revision of the legislation or inspire looking for alternative indicators that can be measured in the next policy cycle.

While the types of impacts considered in impact assessments and evaluations should be the same, the way to measure those impacts (i.e. the methods) should be different. When preparing policies, ex-ante impacts can be estimated with a model. These estimates should be

later compared with real (not estimated) data collected in the evaluation. In other words, re-running the same model in the policy preparation and policy evaluation phase using the same input data would not give evidence on the realisation of impacts.

3.6. Different types of data

The overall strength of evidence comes from the quality of data gathered in previous parts of the policy cycle, from the appropriateness of the methodology applied, and from the concurrence of different parts of the evidence (see Tool #4 (*Evidence-informed policymaking*)).

Scientific evidence on the quality and impact of public interventions, policies and programmes is based on the application of the scientific method to data coming from either:

- observational data, i.e. data collected by passive observation of individuals, firms, communities potentially affected by the policy;
- opinion data;
- randomised control trials (RCT), i.e. data from a controlled experiment.

Assume a new market regulation is introduced, and this is the policy of interest for an evaluation. Data is generated by the functioning of the market to which the new regulation is applied. The data generated by the market is an example of **observational data,** which is typically collected by appropriate reporting provisions in the legal act.

The new market regulation may have affected consumers and firms; both of them are stakeholders, and they may be asked to state their perceptions on the effects of the policy. The data collected on the opinions or perceptions of stakeholders is **opinion data**. Data on stakeholders' opinions is typically gathered through stakeholder consultations (see Chapter 7).

Remark on data types: data on opinions and observational data are *complementary sources* of information. They can be used to provide different types of evidence for the evaluation. **Hence, both types of data need to be collected.**

Observational data can be recorded for single citizens (individuals), products, firms, or geographical areas like regions. For instance, one may wish to record the price increase of a certain type of product sold in a given market, like the price of 1 kg of a certain cheese, protected by Geographical Indication (GI). These types of data are called **microdata** because they are disaggregated (micro), and it would not make sense to record data at a lower level of aggregation (i.e. at a higher level of granularity).

It is possible to aggregate microdata at different levels. For instance, the price of a category of goods can be aggregated; one may wish to record the price of 1 kg for all types of cheese in a given market. Prices can be aggregated across categories of goods and services, and/or at sector level, at national level and so on. These types of data are called **macrodata.** A typical example of an aggregate would be households' consumption of dairy products.

When evaluating policies, it is preferable to use microdata, as it is usually simpler and more convincing to link them to the policy effects. If microdata are not available or accessible, macrodata can also be used, albeit the attribution of the effects to the policy at hand might be more difficult or less direct than when using microdata. For the GI cheese, one may be

interested to observe the time evolution of the GI-protected cheese (the 'treated' group) and to compare this with the time evolution of non-GI-protected cheese (the 'control group'). The time evolution of the aggregate price of cheese may conceal these differences, hence making it difficult to attribute price changes to the GI scheme.

Other data can also be acquired or accessed from market surveillance organisms and regulators, government administrations and other public institutions; this type of data is called registry data or **administrative data**. Administrative data needs to be accessed with due personal privacy safeguards, compliant with the GDPR[915]. Using administrative data is very cost-effective, as the data has already been collected for other purposes (so the acquisition cost is marginal), with large number of individual records and good representativeness (see below for definition).

Government authorities may be unwilling to share data, unless specific provisions to collect and share data exist in the legislation. It is therefore important to mandate their collection and sharing when defining monitoring and evaluation provisions in the policy development stage (see also Section 4). EU decentralised agencies and other EU bodies, managing and regulatory authorities in Member States usually collect data on products, services, individuals, firms that are influenced or affected by EU policy. They are another potential source of data to consider when defining monitoring and evaluation provisions.

Other observational data can be collected via surveys. Survey data may inform both about opinions and perceptions of respondents, as well as on objective data, such as personal income. Self-reporting of such variables in surveys may be, however, prone to various biases and requires triangulation (see section 3.4 above). Overall, the cost of data from surveys is usually higher than the one of administrative data. Moreover, the level of non-response in surveys can be very high.

Remark on the cost of data: administrative data (including data from government authorities and regulatory national authorities) is quite inexpensive as it has already been collected for other purposes when compared to survey data. Moreover, administrative data usually covers the whole population, and hence the sample size is much larger than what can typically be covered by surveys. Conversely, some information can only be collected with especially designed surveys. Hence, **the best approach is to integrate administrative and survey data,** foreseeing which data to link from various sources.

Remark on available secondary sources: The gathering and identification of suitable data often comes down to a meticulous literature review and thorough desk research. The Find-eR search tool gives access to the Commission library collection of e-journals, eBooks, portals and database.

3.7. Data linkage

Different sources of data can be linked together. For instance, the tax records may contain the turnover of a firm and the social security files may contain the number of employees of the same firm. In order to merge or *link* the two sources of information, the same identifier of the

[915] The General Data Protection Regulation (GDPR) from 2018 provides the legal rules to be obliged to in the European Union. See also Tool #55 (*Horizontal matters – publication of responses, data protection, access to documents and transparency register*) on i.a. data protection.

firm needs to appear in both data sets. This process of merging data is called 'data linkage'[916].

Some EU Member States have special agencies whose task is to link data for policy evaluation research. These agencies are usually public and may charge a relatively small fee to perform data linkage.

Remark on data linkage: data from different sources (including government authorities, Regulatory National Authorities etc.) can be linked using unique **identifiers**. For instance, data from different sources can be linked via geographical reference, such as when data refers to the same NUTS3 region. Data on firms can be linked via their name and address or national TAX code, etc. If needed, data can subsequently be anonymized in order shield sensitive information from analysts.

3.8. How should data be collected?

Appropriate methodology needs to be applied to analyse different types of data, to extract evidence on the policy intervention. The correct method is the one that fits the right data, and the process with which this data was collected, as discussed above.

There are methods that are more robust in capturing the causal links between policies and their effects. These methods are called *causal*, and they described in Tool #68 (*Methods for evaluating causal effects*). The ultimate choice of the method will depend on the available data.

Remark on the sequence of questions to ask: One could start with the question: what is the best data-and-method combination for this policy intervention and the evaluation of it?

One can start from listing data, and next, given each dataset, define which method can be best used to produce evidence. Some other times e.g. in structural modelling one can proceed in the reverse order, selecting methods first and data next.

Both opinion data and observational data need to be collected. In both cases, one may need a **sampling design**, i.e. a plan about how to sample from the given population to produce representative samples. Sometimes this plan has already been collected by someone else, for instance by some other agency or institution; in this case one needs to give account of this and to consider this sampling design. Some other times, one needs to directly collect data, and hence one needs to clarify the sampling design.

In business research, companies must often generate samples of customers, clients, employees, and so forth to gather their opinions. Sample design is also a critical component of marketing research and employee research for many organizations. During sample design, the officers responsible for the design must answer questions such as:

- What is the relevant population, sampling frame, and sampling unit?
- What is the appropriate margin of error that should be achieved?
- How should sampling error and non-sampling error be assessed and balanced?

[916] See e.g. Chapter 6 in Crato & Paruolo (2019).

These issues require very careful consideration. Some introduction and reference are given next; these considerations apply both to stakeholder survey as well as to other surveys used to collect observational data.

A few of the above suggestions are summarised in the following box:

Suggestions:

a) Focus on a few outcomes of interest that are close to the objectives in the intervention logic.

b) If impacts cannot be measured ex post it is often not useful to consider them ex ante.

c) Make an ex-ante plan on how to collect data for the ex post evaluation.

d) Try to access data that has already been collected (administrative data, data from government authorities and national regulatory authorities).

e) If you have data in-house (like the ones on inputs), share it with all of the Commission.

f) If you outsource an impact assessment or evaluation to contractors, ask them to hand in the data that they collect.

g) Where possible, survey data should be representative for the population (collected randomly).

h) Report survey data by subpopulations.

i) Report uncertainty estimates from surveys.

4. ADDITIONAL INFORMATION

- An introductory 2 minute video on administrative data is available here.
- An introductory 2 minute video on data for evaluation is available here
- A tutorial (25 minutes) on pros and cons of different types of data, data collection, data access and data merging in relation to causal methods is available here.

Relevant publications:

- Cochran W. (1977) Sampling Techniques, Third Edition
- Groves et al (2009) Survey Methodology Wiley

TOOL #68. METHODS FOR EVALUATING CAUSAL EFFECTS

This tool introduces the principles of ex-post quantification of impacts in evaluations, especially relevant for (i) efficiency (ii) effectiveness (iii) EU value added. **Quantitative analysis** complements **qualitative analysis**: both are essential to build the narrative of the evaluation and to explain the analysis of impacts.

1. CAUSAL EFFECTS

Evaluations aims to answer questions such as: "What were the effects of the policy? Who was affected? What change did it make?" These questions are summarised as "What worked and for whom?" in the following.

The **causal effect** of a policy or intervention is **the consequence (i.e. the effect) attributable to it**. Causal analysis aims to identify the effect of the intervention (called '*x*') on the **outcomes** of interest (both intended and unintended – called '*y*'), considering (or controlling for) other causes of the phenomenon generating the outcomes, including other policy interventions (**contextual indicators**, called '*z*').

Evaluation aims to **draw conclusions about the causal effects of** an intervention, using but going beyond the description of ***what*** happened, and looking for reasons ***why*** it happened (the mechanism through which the causes act on it), i.e., it aims to attribute observed changes to the intervention. In other terms it would like to conclude that "*x* causes *y* controlling for *z*". (i.e. taking 'z' into account)

As an example, one may observe neutral or negative changes in the outcome of interest; one should not (only on the basis of this) conclude that the intervention was ineffective or detrimental. This is because "what would have happened without the policy" called the **counterfactual**, could have been worse. Of course, the same applies to positive changes. Isolating the observed changes and attributing them to the intervention is a challenging exercise but one should **strive to identify causal effects**, if possible (in terms of data and methods).

2. CORRELATION AND CAUSALITY

A simple concept in quantification is the one of correlation. Assuming for a moment that there are no contextual indicators '*z*' to consider, one could wish to measure the correlation between outcome '*y*' and the intervention '*x*', indicated here as corr(*x*,*y*). For the purpose of this section, also assume that "*x* does not cause *y*" (or vice versa) is translated into "*x* and *y* are independent."

This section illustrates that correlation and causality are different, but related concepts. Correlation is a measure of linear association between *x* and *y* and does not have any 'direction' (unlike in cause-and-effect statements)[917]. As it is well known, corr(*x*,*y*) can be positive, zero or negative, and takes values between -1 and +1. corr(*x*,*y*) = 0 corresponds to lack of linear association between *x* and *y*.

What is the relation between correlation and causality? Some variables may be correlated because of an underlying common cause. For instance, there could be a positive correlation

[917] corr(*x*,*y*) = corr(*y*,*x*)

between ice-creams sold and number of shark attacks, but they do not cause each other. A more plausible causal explanation is warm weather causing people to go to the beach and increasing the likelihood of both ice creams sold and shark attacks. It is important not to draw causal conclusions on correlation, and whenever possible go beyond correlations to cause-and-effect relation. Correlation is not causation, but what is causation? To use another example, a headache pill can positively correlate with relieved headache, however, causal inference tries to identify the extent to which the relief is attributed to the pill.

Full causality analysis is not always possible, in these instances correlation (regression) analysis is better than nothing, as long as appropriate limitations are accounted for and explained. Of course, if possible, the aim should be to analyse causal effects, for which appropriate methods are detailed in what follows.

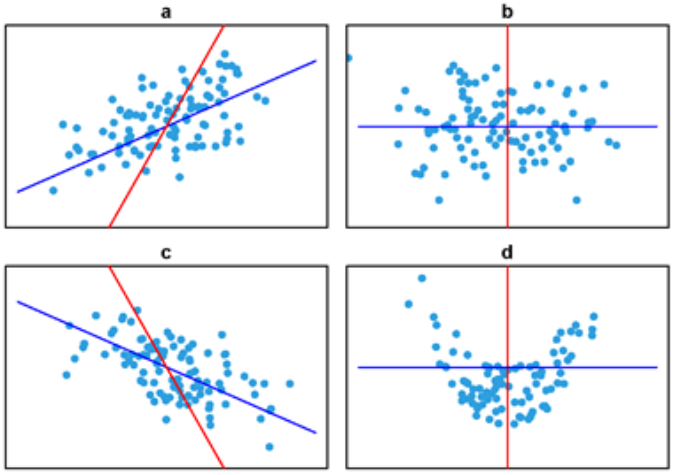


Fig.1. Examples of types of correlations between x (on the horizontal axis) and y (on the vertical axis). The plots contain dots that represent sample values of (x,y) together with regression line of y on x (blue line) and regression of x on y (red line). [1]

- a: positive correlation, corr$(x,y) = 0.7$ (x,y dependent);
- b: zero correlation with independence, corr$(x,y) = 0$ (x, y independent);
- c: negative correlation corr$(x,y) = -0.7$ (x,y dependent);
- d: zero correlation with dependence, corr$(x,y) = 0$ (x, y dependent).

[1] a,b,c: (x,y) jointly normal. d: y generated as $y = 0.25x^2-1+u$, with x normal with mean 0 and variance 4 and u uniform between -1 and 1 and independent from x.

3. COMPARISON GROUP

3.1. A comparison group may help

As explained in Tool #46 (*Designing the evaluation*), one would like to compare policy outcomes with a situation that could have been expected in the absence of a new policy intervention, a 'no-policy-change' scenario[918]. The identification of behaviour in the 'no-policy-change' scenario may use a reliable group of untreated units, called **control group**.

[918] This situation is often called generically 'the counterfactual', which relates to or expresses what has not happened and what could have been.

A control group is made of units that are as similar as possible to the treated ones except from the fact of not being affected by the policy. This control group may be used to estimate the counterfactual scenario of no EU policy.

One needs also to consider the context in which the EU intervention took place, including for instance, national policies with similar objectives (**contextual indicators**) described in Tool #67 (*Data identification for evaluation and impact assessment*). This may allow to 'control' for these national policies when evaluating the EU intervention. The main idea would be to compare the treated group with the control group controlling for contextual indicators.

Lastly, it may not be possible for all EU policies to define a treatment and control group in the strict sense, and several factors and biases should be considered before conducting a causal analysis. Many of them are highlighted below.

3.2. Selection bias, measuring bias and spurious relationships.

In every causal analysis it is important to consider factors that could weaken the causal claim. Some of the more frequent biases that might arise in both observational studies and randomised experiments are selection bias, measurement bias and confounders. These factors should be considered in order to present a robust analysis of causal effects. **The selection bias problem occurs when** the assignment to the treated and the control group was not done at random but out of a selection process related to the outcomes; in this case the population from which the control group was sampled is substantially different from the one of the treated group and comparing them directly would be misleading. For example, if unemployed people self-select into a training programme (the policy intervention), their (hopefully) higher chances to find a job after training may reflect either the benefits of training (i.e. an effect of the policy intervention) or the fact that they may come from the sub-population of unemployed with better motivation and entrepreneurial skills with respect to the ones that did not apply for the training. In the latter case the higher chances to find a job are not an effect of the policy intervention.

This selection bias problem hence complicates the issue of finding a proper control group. Many causal methods solve this issue in various ways. Box 1 provides an example on how to avoid selection bias in student mobility.

Another important factor to consider is whether the causal effect is misgiving or weakened due to the data collection process, **the measurement error bias**. In fact y, x, z could all be measured with an error. Continuing on the example above, training x could be measured with error (for instance some people could have been registered as participants but never showed up at training), their years of work experience z could be measured with error (for instance some years of self-employment may have been not recorded) and their working status 6 months after training y may contain reporting errors (for example people were interviewed 12 months after the end of the training, and some forgot to properly recall their employment status 6 months before). Measurement error (especially in x, z) may induce substantial bias in the estimates of impacts.

A third aspect is related to **the importance of considering contextual indicators**. Contextual indicators refer to variables z that are associated with both the treatment (independent variable x) and the outcome (dependent variable y). When the contextual indicators are not accounted for, one could find a relationship between the independent

variable *x* and the dependent variable *y* when no relation exists. In such a case this could be termed a **spurious relationship**, namely, an observed correlation (or association) without a causal link. For example, the relationship between the beneficial effects of a training programme (dependent variable *y*) and the training program itself (independent variable *x*) does often not existing 'in a vacuum'. Other variables such as the participant's education, income or age could also have a significant causal effect on the outcome. Therefore, contextual indicators *z* are a major threat to the validity of inferences made about cause and effect, when not considered.

Box 1. An example on how to avoid selection bias: student mobility

What is the effect of student mobility (e.g. participation in Erasmus programme) on future employment status?

Mobile students are generally more motivated, have better language skills and hence better labour market outcomes than the average non-mobile students, even before their mobility experience. Hence, **taking any non-mobile students as control group would lead to selection bias**.

Like with many evaluations, different methods can be used for constructing a control group. A good control group would consist of non-mobile students as similar as possible to mobile students in their characteristics.

One possibility would be to create a control group with students who did want to take part in mobility but *just* missed a specific threshold of university qualifications (e.g. a threshold of 50 of a score ranging from 0 to 100) that is necessary to obtain for being eligible for the mobility grant; see e.g. Granato et al. (2020) Comparing these non-mobile students close to the threshold for eligibility (i.e. grades 45 to 49) with those mobile who just made it (i.e. grades 50 to 54) is likely to lead to very similar control and comparison groups, since the **allocation to the comparison and control group around the threshold is likely to be random**. The causal evaluation model adopted in this framework is called **Regression Discontinuity Design** (see below).

Another possibility for constructing a control group is to compare only those non-mobile students with mobile students, who are similar in a large set of characteristics (i.e. upper secondary school mark, motivation etc.) linked to the outcome variable of interest (i.e. being in employment; see Schnepf and d'Hombres 2019). This method of counterfactual impact evaluation is called **Propensity Score Matching**, see (see below).

3.3. Everyone is treated

For regulatory policy, it is often the case that the same rule applies to all citizens or to all firms, so that everyone is treated and there is a lack of a control group. Even in such cases the variation of the treatment could allow for an analysis of causal effects for instance, some methods estimate a **dose-response model**. Namely, those more exposed to the policy intervention are expected to respond more than those less exposed.

3.4. Only aggregate data is available

In case data at the appropriate level of granularity cannot be accessed, but data at aggregate level exists, what can be done? First, one could plan to insert appropriate clauses in the next

policy design phase (ex-ante) so that data at the appropriate level of granularity will be accessible in the future. Secondly, one may work with the **aggregate data** that is available (see Tool #67 for a more detailed discussion on types of data). One approach is to compare estimated models for different periods of time under different policy regimes and compare the structural parameters across periods; this is sometimes called **models for structural change**. Here attribution of the difference to a single policy change is assumed, and this need to be substantiated by appropriate qualitative evidence.

3.5. Can the same policy have different effects at different scales?

The effects of some policies may be different at different scales. Consider for instance some labour market intervention like training, aimed at supporting the long-term unemployed. This intervention may generate higher exit rates from long-term unemployment on a micro scale, i.e. the long-term unemployed subject to the policy intervention may re-enter employment more often or more quickly than the ones not taking training.

At the more aggregate general unemployment rate, this policy may have no effect, as the trained long-term unemployed may have found jobs that have displaced other individuals, with a zero aggregate effect. These more aggregate effects are sometimes called **(general) equilibrium effects**. A consequence of this is that an evaluation needs to decide at what level(s) of granularity it needs to be applied. Different levels may lead to different answers.

3.6. Specific contribution in a policy package

Often, one policy intervention is part of a policy package with general objectives. This is a central aspect for **fitness checks** but applies more generally to most evaluations. For instance, labelling of sustainable products may be part of the European Green Deal, and possibly contributing to other policy priorities. This implies that each evaluation should consider how to address the following problems:

(i) What is the contribution of the single policy intervention to the overall objectives and outcomes of the policy package?

(ii) What is the specific effect of the single policy intervention within the overall policy package, i.e., what is the effect of the single policy intervention that could not have been achieved by the rest of the package?

(iii) Were there trade-offs or synergies with other parts of the policy package?

(iv) Were there trade-offs or synergies with other policy priorities?

Answering these questions requires advance joint planning from the Commission services working on the different related initiatives.

Box 2. Specificities, trade-offs and synergies in product labelling

Consider for instance a regulation overseeing labelling of sustainable products in a market, with the objective to increase sustainability of the economy and to contribute to the overall climate policies of the EU. The specific effects of this regulation may be best measured in the product market where this labelling takes place: possible indicators capturing the specific contribution could hence be the volume and price of goods traded with this label, relative to other unlabelled products with similar characteristics.

A possible trade-off for this labelling regulation could be associated to excessive regulatory burden: if the cost for firms to label their product as sustainable was disproportionally high, this could make the products too expensive relative to competing products, and the overall effect could fail to increase sustainability of the economy overall.

Possible synergies in this area could come from certification of the firm as being sustainable or 'green' overall via related regulation (such as the Green Taxonomy): the effects of the labelling of a product and the certification of the producing firm as 'green' could reinforce each other.

As the example in Box 2 shows, one would need to combine data from product markets and certification of firms to see the effects of these two pieces of legislation in action on their own or in combination.

Causal evaluation methods that address these questions require to have data on inputs and outcomes of different interventions of the policy package and of other policy priorities, hopefully for granular level of data. Moreover, it requires to have contextual information on other national policies that are present in the same policy domain.

4. METHODS

The benchmark for scientific quantitative methods is **randomised control trials,** where units are assigned to treatment or to the control group at random. This solves the selection bias issue. When **randomised control trials** are not possible, one can resort to **quasi-experimental** methods that mimic the experimental design to solve the selection bias, under appropriate assumptions. Other methods can also be used, that are not delivering causal analysis. Conducting causal methods requires knowledge on statistical methods and the JRC is available to offer advice on when and how to use the methods. However, it is important to know the advantages and disadvantages for applying the variety of methods outlined below.

4.1. Randomised control trial (RCT) or experimental approach

Units of the same population are assigned to treatment and control groups at random – e.g. using a lottery; this is called **randomisation.** Randomisation could be used e.g. in the case of over-subscription, when there are more applicants than available slots for a specific programme. When assignment into treatment is random, the average treatment effect is estimated by comparing the average outcome for the treated and the controls; this is very simple and in ideal situations does not require to use data of contextual indicators.

It could also be used in the case of phase-in interventions, i.e. interventions implemented in several phases, due, for example, to limited capacity to serve all the units at the same time: eventually every unit will be treated, but the time of treatment is randomised. A comparison between the unit treated first and the not-yet treated can show the effect of the policy in the short term, as this is equivalent to random assignment.

Key requirements for randomised control trials: random assignment of individuals/firms to the policy intervention.

4.2. Matching methods

Matching methods select a control group by matching each treated unit with at least one untreated unit that is as similar as possible to the treated one, based on a wide range of

characteristics (such as contextual indicators) covered in the data set available (so-called **observable characteristics**). Therefore, the quality of the control group hinges on the richness of the data available and a large sample size.

The most famous of these methods is the **Propensity Score Matching (PSM)**. In this method the matching is performed on an index, the Propensity Score, which is defined as the probability of self-selecting into the intervention, for given different individual observed characteristics.

Box 3. Example on propensity score matching: the evaluation of a European Social Fund (ESF) funded intervention in Flanders (Belgium), 2015-2018 (Canzian et al 2020)

The "Work Experience for Young Persons" (WIJ) programme is targeted at low-qualified young unemployed with the aim of facilitating their entry to the labour market, through an intensive guiding and counselling activity. Individuals registered at the public employment service were selected and assigned to the treated group using criteria based on age, level of education and previous work experience. However, given financial restrictions not all young individuals registered at employment services were assigned to the intervention, and many of them remained untreated.

The analysis aimed at evaluating the impact of WIJ in terms of probability of being employed or re-entering education for young unemployed, some months after the counselling activity was finished. The evaluation of this intervention was done using PSM since the key data requirements were met. First of all, information was available for a good number of observable characteristics, including relevant demographic variables (gender, nationality, date of birth), level of education and previous work history (including months spent in unemployment, employment or inactivity since graduation, number and type of contracts). Second, conditional on these observable characteristics available, it was possible to estimate well individuals' participation in the programme. Finally, a huge number of untreated units were available, sharing similar characteristics with the treated, so that it was possible to select several control units for each treated.

Key requirements for matching methods:

- availability of information on the observable characteristics influencing the selection into the treatment and influencing the outcomes, for both treated and untreated units;
- observable characteristics should be measured before the assignment into the treatment;
- relatively large sample sizes for treated and – in particular – untreated group.

4.3. Regression Discontinuity Design (RDD)

This approach can be adopted when participation in the intervention is determined by a rule or a **threshold** on the value of a **continuous** variable, called the **forcing** variable or the **running variable**. Examples of forcing variables are income, age, or a test score. Many spending programmes have assignment rules based on thresholds: social benefits based on income, scholarships based on grades, employment programmes based on age, etc.

It is common to distinguish between the case when the assignment rule to treated and untreated is applied with no exceptions, called **sharp RDD**, from the case when exceptions to this strict rule can happen; this latter case is called **fuzzy RDD**. Fuzzy RDD is technically linked to Instrumental variables (IV), see below.

Box 4. Example of RDD: the impact of longer unemployment benefits on time in unemployment (Lalive, 2008)

Austria implemented in the late 1980s and early 1990s a programme that extended the maximum duration of unemployment benefits from 30 weeks to 209 weeks of unemployed benefit duration to individuals aged 50 years or older who have been living in certain parts of Austria. Given that the programme was strictly targeted with respect to age and region, a sharp RDD on age and distance to regional boarder is a natural choice to evaluate this policy.

Sharp discontinuities in treatment assignment at age 50 and at the border between eligible regions and control regions identify the effect of extended benefits on unemployment duration.

Key requirements for RDD:

- The probability of participation in the intervention needs to jump at a threshold value of a continuous score (i.e. age, income etc.).
- Enough units (i.e. individuals) should be observed on both side of the threshold (treated and untreated).

4.4. Differences in Differences (DiD)

The method 'differences in differences' compares the changes in the outcome variable over time between treated and untreated units. Provided that (i) in the pre-intervention period, the outcome variables in the treated and control groups follow similar trends (the '**parallel trends' assumption**), and that (ii) the factors explaining differences in outcomes between the two groups **are constant over time**, differences in the trajectory of the outcome variable between the two groups can be attributed to the participation in the programme only.

This approach requires repeated observations over time on both groups (so-called longitudinal data or repeated cross sections).

Box 5. Example of DiD: the Airport Charges Directive (Conti et al 2019)

The Airport Charges Directive (ACD, Directive 2009/12/EC) was approved in 2009 and converted into national legislation between 2011 and 2014, with most Member States transposing in 2011 and 2012. The Directive is meant to be applied only to airports with more than 5 million passengers (or to the country's largest airport if no airports pass this threshold) and involves only a subset of airport charges (i.e. *in scope* charges). The primary objective is to increase the bargaining power of airlines vis-à-vis airports.

One approach to define the treated/control group is to use the population of airports with more/less than 5 million passengers. In this case there are too few airports on each side of the threshold, so RDD could not be used. Hence, this regulatory framework lends itself to a DiD

approach, exploiting the magnitude of the changes in the average level of *in-scope charges* (outcome variable) in treated and non-treated airports before and after the adoption of the directive.

Key requirements for DiD:

- The policy, targeting only a group of units (the treated), should become effective at a certain point in time; this point in time should be known.
- Data for the outcomes of interest should be available both before and after the policy adoption, for both the treated and control group.
- In the pre-treatment period, trends of the outcome variable in the treated group and in the control group should be parallel. This may be checked if a longer time series of the outcome variable before the policy is implemented is available.

4.5. Instrumental Variables (IV)

This method uses an observable variable (called the instrument) which influences the treatment assignment of units but that it is otherwise unrelated to the outcome of interest; this latter requirement is known as the **exclusion restriction**. For example, if a policy offers an unemployment programme at a specific service provider, the distance from each individuals' home to the place where the programme takes place is likely to impact on the participation in the programme but is generally unlikely to be associated with the outcome of the programme.

Box 6. Example of an instrumental variable approach: the Covenant of Mayors (Martelli et al 2018)

The Covenant of Mayors (CoM) was the mainstream European movement for local authorities voluntarily committing to meet and exceed the EU's 20% Green House Gas (GHG) emission reduction target by 2020. In Italian municipalities, the CoM is the best known and recognised action primarily targeting the reduction of GHG emissions.

Martelli et al (2018) want to answer the question "Does the CoM increase the approval of a candidate?" and hence choose as outcome variable the percentage of votes a Mayor receives in re-elections. Treatment status is participation in the CoM, which is not random, and may be associated with positive additional characteristics of the candidate.

The instrument used here is the presence of a so-called 'Covenant Territorial Coordinator,' which provides technical support for drawing up a sustainability plan necessary to join the CoM. This instrument is assumed to affect the probability of treatment (i.e. joining the CoM), but to be unrelated to the political outcome otherwise (exclusion restriction).

Key requirements for instrumental variable approach:

- An instrumental variable is available; an instrumental variable is a variable that affects the outcome of interest only via its influence on the probability to be treated.
- Data on the instrument should be available for both treated and non-treated units.

4.6. What if everyone is treated? DiD with staggered implementation

When every unit is treated, one could use the following scheme. Sometimes units, e.g. individuals, firms, or Member States, adopt the policy or treatment of interest at a particular point in time, and then remain exposed to this treatment at all times afterwards; this situation is called 'staggered implementation'. All units are eventually treated, but not all and every unit is treated at each point of time.

One approach to exploit this situation is to use a DiD approach. The idea is to compare the changes in the outcome of interest for the treated and the control groups. At each point of time, there are units that have already started to be treated, which form the treated group, and other that are waiting to be treated, and they form the control group.

Box 7: Evaluation of the late payment directive (Conti et al. 2018)

Consider the introduction of a directive such as the Late Payment Directive (LPD) aimed at reducing the payment duration between the Public Administration (PA) and the private sector. Assume that the time of implementation varies across Member States, that is a given industry can be treated in a given Member States, but not treated yet in another one. Hence, the impact of the Directive is estimated by comparing changes in the outcome variable (i.e., firms' exit rate, employment rate, investment rate) in the industries in adopting countries to changes in the same outcome variable in the same industries in non-adopting countries, before and after the specific country's enactment of the directive.

Key requirements for staggered implementation with DiD:

- Staggered time of implementation/adoption of the intervention.
- Same requirements as for DiD.

4.7. DiD with intensity of treatment

There are situations where all units comply with the policy intervention (regulatory change), but they differ in the extent to which they are exposed to the intervention. This creates different treatment **intensities,** which can be exploited to identify the effect of the intervention. This approach compares the changes in the outcome variable over time between units more exposed as compared to less exposed ones, using a DiD comparison before and after the intervention.

Box 8: Evaluation of the late payment directive (continued)

Consider again the introduction of the Late Payment Directive (LPD) aiming to reduce the payment duration between the Public Administration (PA) and the private sector. While all industries are subject to the Directive, they may differ in the level of treatment intensity, that is the industries' exposure to the Directive is defined on the basis of the degree of economic dependence of an industry on the Public Administration (PA).

The more an industry sells to the PA (as a fraction of its total sales), the more that industry is exposed to the LPD and hence the higher the expected impact on the firms in that sector. Therefore, the average change over time of the probability to exit the market for a group of

industries more dependent on the PA is compared to the average change of the same variable for a group of industries whose business activity is less and/or not connected to the PA, before and after the adoption of the Directive. Again, see Conti et al. (2018).

Key requirements for DiD with interaction with intensity:

- The possibility to define a degree of exposure to the change in all units of the analysis. The degree of exposure calls for a definition of 'more' treated units as compared to 'less' treated ones.
- Same requirements as for DiD.

4.8. Generalised propensity score (GPS)

The propensity score matching (PSM) method described above matches treated units with similar units that are not treated. The Generalised Propensity Score methods (GPSM) extend this principle, for example, to cases where: (i) different units may be subject to distinct types of treatment; (ii) the units may be subject to the same treatment but with distinct levels or intensities (as in the example above). Under similar requirements as the PSM, the relevant treatment effect compares the impact of different treatments, or the impact of different treatment intensities.

Box 9. Regional growth (Becker et al 2012)

One example of this in the context of the evaluation of regional policies can be found in Becker et al (2012), using data at the NUTS3 level from the last two EU budgetary periods (1994–1999 and 2000–2006) and generalised propensity score estimation to analyse to which extent the goal of fostering growth in the target regions was achieved with the funds provided and whether or not more transfers generated stronger growth effects.

Key requirements for GPSM:

- Availability of information on the observable characteristics influencing the selection into the different treatments or distinct levels of treatment, and the outcome, for all the considered groups.
- These observable characteristics should be measured before the assignment into the various treatments or level of treatment.

4.9. Synthetic controls

The synthetic control method differs from other counterfactual methods because it can be used to evaluate the effect of policies at an aggregate level. For example, a whole country is affected by some specific policy or condition, and one wishes to measure its effect. Assume only one unit is treated, and a few are untreated units. The control group is built as a weighted average of the available set of untreated units, which is called the 'synthetic control' or 'artificial control group' and serves as the counterfactual for the treated unit.

This technique has two main assumptions: first, the selected predictors of the outcome should include variables that approximate the evolution of the treated units without anticipating the effects of the intervention. Second, untreated units used to generate the synthetic control should not be affected by the programme.

Box 10. Example of a synthetic control method (Campos et al. 2019)

Campos et al. (2019) is a recent example of synthetic controls. It aims to estimate the economic benefits of European integration. The key economic and policy question is "What would have been the levels of per capita GDP or productivity in a given country if it had not become a full member of the European Union?" Using synthetic control methods, the authors estimated how GDP per capita and labour productivity would have behaved for the countries that joined the European Union (EU) in the 1973, 1980s, 1995 and 2004 enlargements, if those countries had not joined the EU.

Key requirements for synthetic controls:

- Data on outcome of interest and other controls on one treated and several untreated units over time, before and after the policy.
- Data must be available for several periods before the intervention/policy in the treated unit and the pool untreated units.

4.10. Structural-change time series models

If aggregate data exists on the outcomes of interest and its determinants, one could adopt a 'structural-change' approach for time series models. The idea is to estimate a (set of) relationship(s) in a time series model before and after the policy intervention, and then to compare the changes in parameters associated with these two different periods. This would signal evidence of changes in the relationship associated with the policy change. The attribution of these effects to the single policy intervention is the part of this approach that is most difficult to establish.

This approach is common in monetary economics for international real interest rates and inflation, for the equity premium, global house prices, CO_2 emissions, etc.

Box 11. The effects of the euro on national Phillips curves (Girardi and Paruolo 2013)

An example of application to the effects of the euro on the Phillips curves in the EU is Girardi and Paruolo (2013), who investigate possible structural changes induced by the euro on the macroeconomic relations among wages, prices and unemployment for the five major European economies. The dynamic adjustment and the level relations are found to be different across subperiods as well as across countries; for an interpretation of results see the original paper.

Key requirements for structural-change time series models:

- Data on outcome of interests and their determinants across time, before and after the policy.
- Enough data should be available for both periods to allow separate estimation.

4.11. Correlation (or regression) models

If none of the above methods can be applied, it is usually possible to run regression (or correlation) analysis among the outcome variables of interest and other determinants. A large degree of covariance between the treatment and the outcome, controlling for all contextual indicators, is suggestive of the presence of causality, although it may not necessarily prove it.

Many of the causal methods reported above are based on some form of regression; what makes these causal methods attractive is that they state explicitly under what conditions the obtained estimates have a causal interpretation. Whenever possible, it is advisable to provide the list of assumptions that would make the estimates causal and discuss why it is expected (or not expected) that they are satisfied in the specific application.

4.12. Mixed methods for evaluation

Quantitative methods, like the ones presented in the present tool, answer the question "Was there a causal effect of the program or regulation, when did it occur and what is its size?" Once the effects have been estimated, qualitative methods can provide useful information to understand the mechanisms at work behind the estimated impact, also especially when the effects are estimated to be null. This mixture of qualitative and quantitative methods is called the **mixed method approach**.

Qualitative methods include, for example, focus groups or in-depth interviews with selected beneficiaries; these can provide the context and help explain the obtained quantitative results, getting insight on the mechanisms underlying the estimated effect of the intervention. They can also provide information on the implementation process, which can be different from how the intervention was initially planned.

Any lack of success of the policy intervention could be due either to design failure, when the proposed policy design was not appropriate to achieve the intended objectives in the particular context or implementation failure, when the project was not implemented as originally planned. These considerations are important to guide re-design of future policies.

Mixed methods incorporate qualitative contextual analysis examining the influence of external factors (via the use of participant interviews or focus groups) with process analysis (via similar qualitative techniques applied to the programs' organisational processes) to assess the process of project implementation and how this affected program outcomes and the estimated impact.

Some of these methods can offer insights while the intervention is still ongoing, i.e. during the monitoring phase, and they may help to re-target the intervention or to improve some of its features, even before impacts are estimated using quantitative methods.

Qualitative techniques at the beginning of the intervention plan can help to target specific groups of the population, which are not registered in official administrative data and may

therefore be not easily reachable through regular surveys' techniques. For example, some marginalised social groups, such as the homeless or irregular migrants cannot be found in registry data, do not appear in official statistics; they can probably be only reached via in-depth field interviews. Similarly, information about some sensitive topics (for example, domestic violence) could be obtained more easily with this kind of interviews, rather than with standard survey data or data from official registries.

Qualitative data and methods have been used extensively to complement quantitative analysis of the effect of interventions in developing countries, Bamberger et al. (2010).

Box 12. Qualitative methods as complements ESF evaluation for Flanders (Belgium) (Canzian et al 2020),

This intervention was evaluated using matching techniques, which compared treated individuals to similar non-treated ones. The results show that people participating in this intervention did not show higher employment probabilities than similar people not participating. However, very little was known about the selection process (how case-workers decided which individuals would be assigned to the intervention, once the age and education requirements were met) and on the real contents of the interventions.

Qualitative interviews with the case workers can help shed more lights on the selection process. In addition, focus groups or interviews with some of the participants, or qualitative interviews with some of the NGOs in charge of providing the support could provide insights on why the intervention did not bring the expected results. However, qualitative data alone would not have answered the question "Has this intervention worked?"

5. ADDITIONAL INFORMATION

- A tutorial (25 minutes) on how to choose the right causal evaluation method, given policy's features is available here.
- An introductory 5-minute video on RCT and Matching is available here.
- An introductory 2-minute video on RDD is available here.
- An introductory 2-minute video on DiD is available here.
- An introductory 2-minute video on IV is available here.

TOOL #69. EMERGING METHODS AND POLICY INSTRUMENTS

This tool discusses the regulatory sandboxes and behavioural insights, two emerging approaches that may get a bigger role in the policy assessment in the future.

1. REGULATORY SANDBOXES

1.1. What is a regulatory sandbox and why is it relevant for policymaking?

Regulatory sandboxes are a relatively new policy instrument. They are part of efforts by regulators across the globe to tackle regulatory challenges generated by technological transformation, and the emergence of new products, services and business models[919]. Although no commonly agreed definition exists, **regulatory sandboxes can be broadly described as schemes that enable firms to test innovations in a controlled real-world environment, under a specific plan developed and monitored by a competent authority**[920]. They are usually organised on a case-by-case basis, include a temporary loosening of applicable rules, and feature safeguards to preserve overarching regulatory objectives, such as safety and consumer protection. Two approaches are theoretically possible to set up a sandbox: one where the request (and identification of a regulatory barrier) is initiated by innovators[921], and another, where the regulator identifies legislative provisions for testing and calls for applications by interested organisations[922]. Additional approaches or a combination of the above may emerge with time.

Their novelty limits the comparability of existing experience; however, current regulatory sandboxes tend to share the following characteristics:

- **Genuine innovation**: the products / services / business models admitted to a sandbox should represent a genuine innovation, not currently available in the market. A new use of an existing technology can also qualify;
- **Societal and/or consumer benefit**: these innovations are expected to deliver consumer and/or wider societal benefits, for instance by addressing unmet social needs or by contributing to policy objectives on e.g. environmental protection, financial stability, competitiveness, and so on;
- **Readiness for testing**: an innovation is advanced enough to be tested in a controlled environment/market and the relevant legislative barrier is identified; theoretical links between an innovative idea and existing rules are not sufficient to set up a sandbox;
- **Defined scope and time**: the boundaries of a regulatory sandbox may be grounded in law (e.g. an experimentation clause)[923]. In any event, boundaries are established ex ante and usually clarify the legislation and sector(s) covered by the test, its duration and exit conditions. This approach ensures legal predictability and facilitates measuring and evaluating sandbox outcomes;

[919] For an overview, see Attrey et al. (2020) and Lesher (2020).
[920] See e.g. Report from the European Supervisory Authorities (ESAs) Joint Committee (2019); Federal Ministry for Economic Affairs and Energy - BMWi (2019), Making Space for Innovation: The Handbook for Regulatory Sandboxes; Office for Gas and Electricity Markets (2018), What is a regulatory sandbox?.
[921] Some examples are available in ESAs Joint Committee (2019: 22-24).
[922] See for instance the German experience as described in BMWi (2019).
[923] On this point see Tool #22 (*Research and innovation*) and the Council conclusions of 16 November 2020.

- **Safeguards**: the purpose of a sandbox is not deregulation. Hence, even in a controlled setting, appropriate safeguards to preserve policy objectives and legal requirements apply (e.g. safety when testing self-driving vehicles).

Box 1: recent examples of regulatory sandboxes in the EU

- **EU-level: Artificial Intelligence Act**

Article 53 of the Commission's proposal provides the general framework for the formal establishment and operation of artificial intelligence (AI) regulatory sandboxes by one or more Member States competent authorities or the European Data Protection Supervisor. The proposal states (Recital 71) that "*competent authorities from one or more Member States should be encouraged to establish artificial intelligence regulatory sandboxes to facilitate the development and testing of innovative AI systems under strict regulatory oversight before these systems are placed on the market or otherwise put into service*". This approach is meant to ensure compliance with the requirements of the AI Regulation and, where relevant, other Union and Member States legislation supervised within the sandbox. Article 54 also provides a special legal basis for the processing of personal data lawfully collected for other purposes that can be re-used for the development − in the sandbox − of certain AI systems in the public interest, subject to additional safeguards and conditions.

- **EU-level: Pilot Regime for Distributed Ledger Technology (DLT) market infrastructures**

The proposed Regulation on DLT market infrastructures aims to enable market participants to operate a DLT market infrastructure (either a DLT multilateral trading facility or a DLT securities settlement system) under certain conditions. It establishes requirements for acquiring a permission to operate a DLT market infrastructure, sets limitations on the transferable securities that can be admitted to trading, and frames the cooperation between the DLT market infrastructure, competent authorities and the European Securities and Markets Authority (ESMA). The proposed regime concerns a limited set of assets and transactions. The overall objective is to remove regulatory hurdles to the issuance, trading and post-trading of financial instruments in crypto-asset form and for regulators to gain experience on the application of DLT in market infrastructures. The Regulation mandates ESMA to carry out a review on the application of the pilot regime three years after its entry into force. It proposes safeguards to ensure consumer protection, market integrity and financial stability. All participants will also have to provide a clear exit strategy, to ensure smooth transitions once the pilot period is over.

- **Germany: transport of medical samples by drone**

The regulatory sandbox Medifly Hamburg allows for the transportation of sample tissue between hospitals located in the same urban area. The sandbox is backed by the Hamburg Authority for Economy, Transport and Innovation, and involves Hamburg's aviation authority and the relevant air traffic control office. The participating consortium, led by the Centre of Applied Aeronautical Research, includes a research institution, software companies, and a drone operator. The sandbox is based on an experimentation clause in Section 21b subsection 3 of the Rules of the Air Regulations. Six test flights were successfully carried out in February 2020.

- **France: facilitating innovative projects for collective self-consumption of electricity**

A derogation to articles L. 315-2 and L 315-3 of the Energy Code aims to facilitate the development of innovative projects in the area of collective self-consumption of electricity. The derogation widens the boundaries of collective self-consumption so that local facilities, larger than those originally permitted under the Energy Code, can fall under the self-consumption definition. The experiment may also remove the 100 kW threshold (article L. 315-3 of the Energy Code) linked to the applicable tariff for the use of public electricity networks. The sandbox runs for five years and is operated under France Experimentation, an initiative by the French Ministry of Economy and Finance.

- **Pan-European blockchain regulatory sandbox**

The EU Member States, Norway and Liechtenstein signed a Declaration creating the European Blockchain Partnership to establish a European Blockchain Services Infrastructure (EBSI) and support the delivery of cross-border digital public services, with the highest standards of security and privacy. In cooperation with the European Commission, the European Blockchain Partnership is now planning a pan-European regulatory sandbox to become operational in 2021/2022. Use cases covered by the sandbox may include data portability, B2B data spaces, smart contracts, and digital identity (Self-Sovereign Identity) in the health, environment, mobility, energy and other key sectors.

Regulatory sandboxes present both advantages and difficulties for all parties involved. Provided the concerned firm(s) can meet the requirements to take part in a sandbox, advantages include the possibility to test own innovations in a real-life setting, and gaining a better understanding of applicable rules, particularly when these fall in the remit of different regulators. Participation in a sandbox may also facilitate access to finance and reduce time-to-market for the innovator.

From a regulator's perspective, sandboxes allow some degree of flexibility without giving up regulatory standards; they facilitate learning, keeping up with developments in the sector, and highlight the implications of existing rules on cross-sectoral innovation and on innovation happening at the 'periphery' of the regulator's competence. They strengthen ties between regulators from different policy fields. Overall, these features can contribute to resilient and relevant legislation. On the downside, regulatory sandboxes may alter the level-playing field in the market; and can increase risks of market fragmentation and 'regulatory arbitrage' if sandboxes for the same rules/innovation lead to different results across the EU. They also require significant resources and time, as well as dedicated skills, that are also needed for 'core' regulatory functions.

Sandboxes are one of the most recent tools of adaptive regulation[924]. Other forms of experimentation are available and may be more appropriate for a specific case, for instance when a clarification of how existing legislation applies to an innovation can be provided through interpretive guidelines and without additional testing[925]. In fact, sandboxes may be the follow-up to other, looser forms of experimentation, if these did not yield the desired clarity on the link between an innovation and the existing regulatory framework.

[924] For further details on anticipating change and ensuring that policies and legislation are future proof, see Tool #20 (*Strategic foresight for impact assessments and evaluations*)

[925] See Tool #22 (*Research and innovation*) for other forms of experimentation. For additional examples, see ESAs Joint Committee (2019) on innovation hubs in fintech.

1.2. Elements to consider before setting up a regulatory sandbox

As with other 'better regulation' tools, the principle of proportionate analysis applies when considering whether to set up a regulatory sandbox. A valuable starting point would be to draw a list of existing experimentation tools in the policy field under consideration, including examples at the national level. Such a stocktaking exercise can already shed light on potential frictions between legislation and selected innovations. It may well be that guidelines would already reduce regulatory uncertainty, without the need for temporary exemptions or testing. At EU level, another potentially relevant source of evidence are innovation deals[926], if any have been concluded in the policy field concerned. Similar initiatives also occur at national level, for instance through innovation hubs.

For further guidance, Table 1 includes a set of questions to consider before deciding whether to establish a regulatory sandbox. The list is divided per focus area and draws on key elements that turned out to be relevant in the limited experience with sandbox implementation. The list is not exhaustive, and practice will complement it in the future.

Note also that existing regulatory sandboxes are limited to specific policy areas (e.g. financial services, energy, digital technologies) and usually implemented locally, as this is where the regulator can more easily control the parameters of the sandbox experiment. One of the main difficulties of a regulatory sandbox is to scale-up the results observed in the testing environment to the wider market. At EU level, an additional challenge is worth mentioning: the impact on the Single Market and the risk of fragmentation if sandboxes for the same innovation are implemented in an uncoordinated manner in different member states. This risk is already known to regulators and various approaches are being considered to mitigate it.

Table 1: Questions for the set-up of a regulatory sandbox

Focus areas	Possible questions	Remarks
Innovation and the market	– Which features of the product/service/business model qualify it as a genuine innovation? – What alternatives to the innovation exist on the market? To what extent are they comparable? – What are the main competitors of the firms concerned? How can fairness be ensured for the firms that do not take part in the sandbox? – Which criteria will be used to establish that an innovation is beneficial?	– This focus area is at risk of subjective judgment and thus open to contestation during e.g. a standstill period before the final list of firms admitted to a sandbox is made public. It is important to ensure transparency on how the relevant criteria are established and applied, and do so early in the process.
Applicable rules and flexibility	– Which body of legislation is relevant for the regulatory sandbox (e.g. banking regulation, data protection, liability for automotive vehicles, etc.)? – Is this a cross-sectoral sandbox, involving different government departments? – Is the regulatory barrier for testing precisely identified? – Who establishes which regulatory requirements/barriers will be covered by the sandbox? Are these chosen by the regulator or	– Existing rules can be customised/made flexible in various ways: e.g. by relaxing or suspending applicable requirements for a limited time and for selected innovations; by waiving enforcement against usually non-permitted behaviour. Another option could be to introduce a limited (in time and scope) but blanket exemption from existing rules: this approach would cover more than one

[926] See Tool #22 (*Research and innovation*).

	identified exclusively by the applicant? – What are the objectives of the relevant legislation that need to be safeguarded during implementation? – What is the scope for making regulatory requirements more flexible? Which form should this flexibility take? – Are there risks of fragmentation for the EU Single Market?	type of innovation thus levelling the playing field. – In case of cross-cutting policy issues, are there different learning/testing objectives for each area of legislation? – Important to clarify that the purpose of the sandbox is not to deregulate.
Access to the sandbox	– How will the selection criteria be outlined in an unambiguous way in the application form? – How is fairness in access ensured for all applicants? What mechanisms are needed to ensure that selection criteria are applied consistently? – Is there a standstill period for unsuccessful applicants to contest the decision leading to their non-admission to the sandbox? – Is there any form of support (e.g. guidance, funding, mentoring) envisaged for applicants?	– Consider the objectives of the sandbox and available resources when preparing the application process: are applications welcome at fixed dates or all the time? Do you have enough resources to react in each case? – Clarify that the sandbox is not an endorsement of the innovation being tested (risk of affecting competition in the market, uneven playing field).
Design and implementation	– What are the goals of the regulatory sandbox? What are its limitations? What indicators will be used to monitor progress and to correct course if needed? – How many companies/innovations can be meaningfully observed in the sandbox? What happens if a company exits the sandbox before the end? – What criteria will be used to close/exit the sandbox? What could be possible consequences on the market, e.g. if a product is discontinued as a consequence of the sandbox?	– How will you clarify what this sandbox is not about?
Evaluation and learning	– What will success look like? – How will you establish if the results of the sandbox can be scaled up, beyond the controlled environment? What risks could materialise when scaling up and how can they be mitigated?	– Ideally, the main evaluation criteria should be established ex ante. If new elements emerge during implementation, they should be integrated into the evaluation strategy in a fair and transparent manner.
Time and resources needed	– Is there any experience at national level, in or outside the EU with a sandbox in this area? If so, can the findings be used as a starting point? – Are sufficient resources available to set-up, run and exit the sandbox? Is there a need to coordinate with other DGs, member states, sectoral regulators and other competent authorities? What are the resources implications of coordination? Are all the parties involved equally equipped to sustain the necessary effort over time?	

1.3. Regulatory sandboxes in 'better regulation'

If available at the early stages of policy preparation, the findings of a regulatory sandbox can be used – together with other sources of evidence – to inform impact assessments and in particular the problem definition and the baseline scenario. Insofar as they provide indications on how a given innovation interacts with applicable legislation, the results of a sandbox may also be used to estimate the impacts of policy options affecting the regulatory environment (e.g. relaxing certain licensing requirements). When doing so, it is important to always consider whether the indications provided by the sandbox remain true when scaling-up. If potential new risks and positive/negative impacts are likely to derive from scaling-up or from an EU-wide application, these should be factored in the analysis.

Regulatory sandboxes may also be useful for an evaluation or fitness check, when specific regulatory barriers to innovation have been signalled during public consultation, through the Fit-for-Future Platform and other channels (e.g. innovation deals). In this case, the regulatory sandbox can inform possible future approaches to tackle these barriers and make the corresponding rules more adaptive and future-proof.

1.4. References and additional sources

- Attrey, A. M. Lesher and C. Lomax (2020), The role of sandboxes in promoting flexibility and innovation in the digital age, OECD Going Digital Toolkit Policy Note N. 2.
- European Supervisory Authorities Joint Committee (2019), FinTech: Regulatory sandboxes and innovation hubs, Report.
- Lesher M. (2020), Bringing new digitally enabled products and services to market: Sandboxes and the role of policy experimentation, Vox EU.
- Parenti R. (2020), Regulatory Sandboxes and Innovation Hubs for Fintech, Study for the committee on Economic and Monetary Affairs, European Parliament.

For further information, please contact: RTD-Innovation-Principle@ec.europa.eu

2. BEHAVIOURAL INSIGHTS

2.1. Main features

Behavioural insights (BI) are **evidence-based conclusions about human behaviour**. They provide a better understanding of how people think, act and feel. Behavioural sciences arrive at these insights through systematic observation and analysis.

Behavioural insights show that human beings are often not rational. They do not always base their decisions on an analysis of all possible courses of actions. **Policy initiatives may fail if they expect rational behaviour by the public.** By understanding how people really behave, we can make policies more effective.

Behavioural insights **can contribute to the EU impact assessment process**. They are not restricted to any particular policy area. Rather, they are relevant when the effectiveness of a policy depends on human behaviour. This will be the case when the policy seeks to change behaviour. But it also applies when the public's reaction to the policy is key to its success.

BI can contribute to the IA process when defining the **problem** or the **policy options**.

- If a policy problem has a behavioural element to it, BI can help identify the source of the problem. E.g., BI can explain why farmers do not adopt greener technologies (even if it is in their interest).
- BI can also help identify policy options. E.g., behavioural experiments tested energy labels for electrical appliances. They found that an 'A-to-G' energy label worked best. EU energy regulation has now incorporated this finding.

The impact assessment process often requires a **quantification of the impact** of a given policy option. Depending on the method used, behavioural insights may or may not be able to provide this. For instance, laboratory experiments can test how a policy option may impact behaviour and establish the direction of the impact. Calculating its size, however, may imply testing in real life conditions.

A good behavioural study **needs time**. In particular, experiments and randomised control trials (RCTs)[927] need at least 5-6 months. First of all, a thorough literature review is needed to gather relevant evidence and develop a sound protocol design. Then, programming, recruitment of subjects, data collection and analysis are additional time-consuming tasks.

For this reason, the potential future needs should be well anticipated, and behavioural insight studies should initiate well before any concrete impact assessment work starts (e.g. in the context of evaluations or as general policy development work for policies which depend crucially on human behaviour). Moreover, behavioural insight studies may well go beyond the policy options that are usually covered in impact assessments, such as detailed issues in the legal drafting.

2.2. Practical implementation steps

The acronym **DO IT** (define, observe, identify, test) helps to remember the practical steps for applying behavioural insights (see Figure 1).

- **Define** the behavioural element.

This first step is very important and will determine whether a behavioural insights approach is relevant and can add value. It implies defining the relevant behaviour and establishing how it relates to the policy problem. For example, do we need to change behaviour or understand it better? In the case of energy labels for household appliances, the Commission was concerned that consumers did not clearly understand the energy efficiency labels. A behavioural study then showed that alphabetic scales worked better than numeric scales, and that an 'A to G' scale worked better than the existing 'A+++ to D' scale. The EU regulation on energy efficiency labelling incorporated this insight accordingly[928]. Another issue to consider is *whose* behaviour is relevant. Behavioural insights usually focus on an individual's behaviour. However, policymakers might also want to understand or change organisations' behaviour. This may imply a different approach of behavioural insights.

[927] See Tool #68 (*Methods for evaluating causal effects*)
[928] https://ec.europa.eu/energy/sites/ener/files/documents/1_EN_ACT_part1_v6.pdf

- **Observe** the behaviour and try to understand it.

Before targeting a behaviour, we need to observe it and try to understand what lies behind it (i.e. what motivates it). For this, we should rely on a literature review or preliminary empirical analysis. How do people think, act and feel in relation to the policy problem? Is there some 'ideal' behaviour that people are not showing in the first place? Why would this be? Do they exhibit some kind of bias? Are other stakeholders exploiting these biases for their own interest? For example, the Commission observed that, in online transactions, pre-checked boxes were leading consumers to make choices that were not in their best interest. A literature review confirmed that these boxes were capitalising on the 'status quo bias', which makes people stick with the default option. The result was a ban on pre-checked boxes, made explicit in art. 22 of the Consumer Rights Directive (2011)[929].

- **Identify** policy options to address this behaviour.

Behavioural insights can help identify policy options to tackle the behaviour that lies at the core of the policy problem. These options can include 'hard' regulation (see energy labelling and pre-checked boxes examples above) or 'soft' behavioural interventions, like *nudges*. Nudges are changes to the environment in which a person makes a decision, aiming to help them make the best decision for themselves. However, they preserve the individual's freedom to choose. For example, a behavioural study tested possible nudges to curb problematic online gambling behaviour. Some of those implemented *in-gamble*, which effectively interrupted human-machine interaction, were effective. Those implemented *pre-gamble*, like self-commitment strategies, were not. This evidence on potential policy remedies was incorporated in the EU Recommendation on online gambling services[930].

- **Test** the effectiveness of these options.

Applying behavioural insights to policymaking is more than bringing a better understanding of behaviour. It is also about testing policy options empirically. These tests are commonly experiments, conducted in a controlled environment, which determine which policy options could be effective. A policy that fails to meet its objectives is an expensive policy. Spending some time and resources testing it before implementing it makes sense. Testing, with whatever method of research, might have some limitations. Yet it is far better to proceed based on some evidence than based on none. For example, the revision of the Tobacco Products Directive put forth a set of pictures and warning messages to be included in tobacco packages to dissuade people from smoking[931]. These were not picked out of a hat, but rather were the result of a series of behavioural experiments that tested the effectiveness of different pictures and messages.

There is no single best method for observing and testing behaviour (for the **O** and the **T** described above). A number of different methodological options are available:

- *Literature reviews* will teach us a great deal about the behaviour in question. Only if this process leaves questions unanswered should we conduct empirical work.

[929] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32011L0083&from=EN
[930] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014H0478&from=EN
[931] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0040&from=EN

- *Qualitative research* includes in-depth interviews, focus groups and participant observation, among others. They can describe the meaning people associate with their actions. This is key to understanding why they behave the way they do.
- *Surveys* measure how people *talk* about behaviour (i.e. what they did or what they intend to do). Results are generalisable but they are not as reliable as measures of actual behaviour.
- *Experiments* observe actual behaviour in a controlled environment. They compare the behaviour of a treatment group, which receives an intervention, and a control group, which does not. If there is a difference, it will be due to the intervention (i.e. experiments can establish causality).
- *Randomised controlled trials (RCTs)* are similar to experiments, but in real-life settings. They do not rely on small samples, but on large sectors of the population. Findings can be quite persuasive but may not be transferable to other contexts.

2.3. Additional information

- Seven Points to Remember when Conducting Behavioural Studies in Support of EU Policy-making (JRC report, 2015)
- Qualitative methodology in behavioural studies for EU policymaking (JRC report, 2016)
- Behavioural Insights Applied to Policymaking (JRC report, 2016)
- Applying Behavioural Science to EU Policy-Making (JRC report, 2013)

Figure 1: Behavioural Insights applied to a policy process