



Climate Policy Ex-Post Evaluation Workshop: Guidance document on climate policy evaluations

Capacity Building to Facilitate Implementation of
Effort Sharing Legislation

23rd September 2020



Agenda



Key characteristics of the guidance for ex-post evaluation



From a bibliography to a reviewed guidance



Design of the guidance? How to use the guidance?



Overview of methodologies



Where to find Member States' practices in the guidance? Case studies



Q & A

Agenda



Key characteristics of the guidance for ex-post evaluation



From a bibliography to a reviewed guidance



Design of the guidance? How to use the guidance?



Overview of methodologies



Where to find Member States' practices in the guidance? Case studies



Q & A



Key characteristics of the guidance for ex-post evaluation of climate policies in field of Effort Sharing

The guidance ...

- ... focuses on **priority evaluation needs** expressed by the Member States;
- ... provides **practical and hands-on** guidelines on key topics;
- ... is concise with **links** to additional material and documents;
- ... is supported by examples and **case studies from bibliography**;
- ... focuses on ex post evaluation only, but multiple evaluation criteria;
- ... is aligned with the **better regulation** guidelines and better regulation toolbox;
- ... aims to improve completeness of reporting under the Energy Union Governance;
- ... should support transparent reporting on evaluation results.



From a bibliography to a reviewed guidance for ex-post evaluation

- From a **bibliography** of 95 ex-post evaluations to a guidance illustrated by good practices
- Special attention to **priority evaluation issues** as formulated by Member States during an online survey (February 2019)
- Review by Member States to increase the added value of support on ex-post evaluation:
 - First review January 2020 of draft, partially completed guidance
 - Second review of completed guidance by Member States during Summer 2020





Bibliography of Member States' ex-post evaluation practices

- Repository of 95 studies on ex-post evaluations across sectors relevant to Effort Sharing
- Coverage: sector – policy instrument / evaluation criteria / data availability

Per sector - per instrument type

	Scope - Buildings	Scope - Transport	Scope - Agriculture	Scope - industry	Scope - waste	Scope - cross-cutting	TOTAL
Economic	18	12	9	10	3	10	27
Fiscal	29	16	6	17	3	15	42
Voluntary/neg	9	6	4	7	2	5	16
Regulatory	25	17	8	14	2	16	36
Information	16	7	6	7	3	3	19
Education	4	1	1	1	1	-	4
Research	1	-	1	-	-	1	3
Planning	2	4	2	1	1	1	7
N/A	-	-	-	-	-	-	0
Other	1	1	1	1	-	-	0
TOTAL	61	33	17	33	7	29	154



Bibliography of Member States' ex-post evaluation practices

- Repository (.xlsx) of 95 studies of ex-post evaluations across sectors relevant to Effort Sharing
- Multiple data fields and categories:
 - Identification of study
 - Scope of evaluation
 - Quantification methodology
 - Indicators or factors considered in monitoring & evaluation process
 - Quantitative results (emissions and costs)
 - Evaluation criteria
 - Replicability to other Member States

	A	B	C	D	K	AF	AL	AZ	BF	BI
1				Identification	Evaluation General Approach/Scope	Quantification Method & indicators- Details	Quantification results following MMR reporting	Evaluation Criteria	Replicability	Priority Evaluation Issues?
2	Evaluation Title	Name of policy or measure	ID							
3	Evaluatie Meerjarenafspraak Energie Efficiëntie 2008-2020 (MJA3)	Evaluatie Meerjarenafspraak Energie Efficiëntie 2008-2020 (long-term agreement on energy efficiency)	1							
	RHI Evaluation: Synthesis	Renewable Heat Incentive	2							

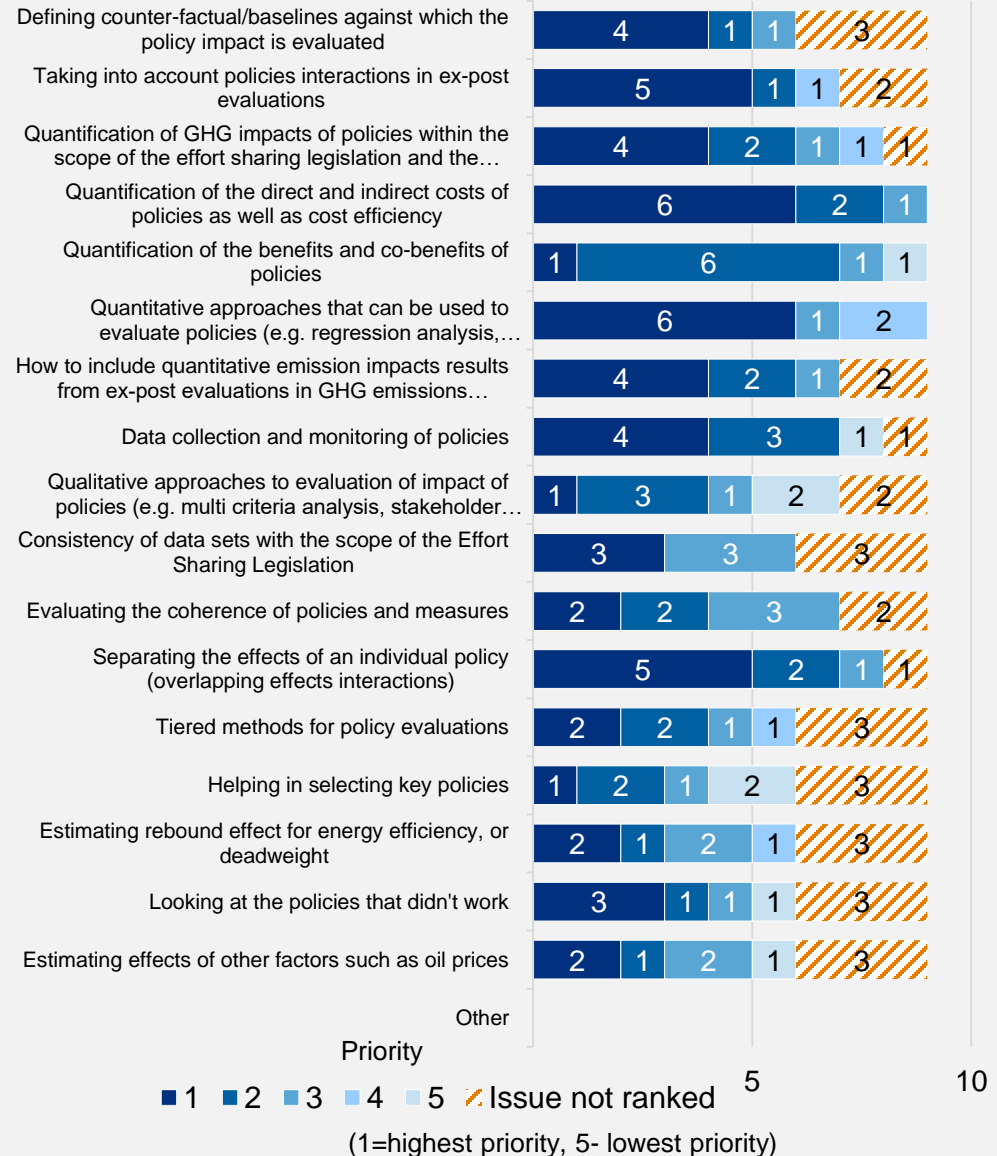


Identification of priority evaluation issues

- Collection of evidence or data for ex-post evaluations;
- Quantitative methodologies that can be used to evaluate policies, including cost effectiveness and cost benefit analysis of emission reductions;
- Key evaluation issues or methodological challenges, such as:
 - Policies interactions in ex-post evaluations;
 - How to include quantitative emission impact results from ex-post evaluations in GHG emissions projections;
 - Quantification of GHG impacts of policies within the scope of the Effort Sharing legislation and the impacts outside of the scope (ETS).

Guidelines about collection of evidence, quantitative evaluation approaches (incl. cost effectiveness) and other evaluation issues.

Source: Response to survey of Member States on Capacity Building Needs and Priorities
 "Interest from Member States in receiving further guidance on specific evaluation issues."



Agenda



Key characteristics of the guidance for ex-post evaluation



From a bibliography to a reviewed guidance



Design of the guidance? How to use the guidance?



Overview of methodologies



Where to find Member States' practices in the guidance? Case studies



Q & A

Design of the guidance

- Step-wise, concise guidance
- Fixed structure per topic
 - What?
 - When?
 - How?
 - Tools & data sources
 - Related topics
 - Links
- Clear pointers to other documents
- Existing case studies to illustrate methodology in practice

Table of contents

1	Introduction	1
2	What is an evaluation?	5
2.1	Introduction.....	5
2.2	Frameworks for policy evaluation.....	11
2.2.1	Intervention logic model.....	11
2.2.2	Counterfactual analysis.....	14
3	Toolbox for ex-post evaluation: methodologies for collecting evidence	17
3.1	Summary.....	17
3.2	Surveys.....	19
3.3	Systematic literature review.....	27
3.4	Focus groups.....	31
3.5	Interviews.....	33
3.6	Monitoring performance data and new data collection.....	37
4	Toolbox of ex-post evaluation: analytical methodologies	46
4.1	Summary.....	46
4.2	Indicator analysis.....	48
4.3	Cost effectiveness analysis and cost benefit analysis.....	56
4.4	Regression analysis.....	68
4.5	Decomposition analysis.....	73
4.6	Multi-criteria analysis.....	81
5	Key evaluation issues and methodological challenges	85
5.1	Assessing policy interactions.....	85
5.2	Rebound effect.....	89
5.3	Free rider effect.....	95
5.4	Uncertainty.....	98
5.5	Emission factors.....	101
5.6	Assessing co-benefits.....	104
5.7	Consistency between ex post information and projections.....	109
5.8	Splitting impacts between Effort Sharing and ETS.....	112
6	Case studies	117
	Case #1: Application of Surveys in Agriculture Policy Evaluations (United Kingdom and EU).....	118
	Case #2: The impact of Swedish SO ₂ policy instruments on SO ₂ emissions 1990-2012 (Åström et al., 2017).....	124
	Case #3: Applying ex-post index decomposition analysis to final energy consumption for evaluating European energy efficiency policies and targets (Reuter et al., 2019).....	128
	Case #4: Evaluation of Dutch long-term agreement on energy efficiency (Meerjarenspraak – MJA3) 2008 – 2020 (Ecorys, 2013).....	133
	Case #5: Dutch Energy Investment allowance (EIA) 2012-2017 (CE Delft, 2018).....	137
	Case #6: Economic evaluation of French écopastille (eco-tax bonus-malus & super-bonus new vehicles) 2008-2012 (CGDD, 2013).....	141
	Case #7: Free riding on tax credits CIDD for home insulation in France (Nauleau, 2014).....	145
	Case #8: Evaluating public policy instruments in the Greek building sector (Spyridaki et al., 2016).....	149



How to use the guidance?

- Step-wise, concise guidance
- Fixed structure per topic
 - What?
 - When?
 - How?
 - Tools & data sources
 - Related topics
 - Links
- Clear pointers to other documents
- Existing case studies to illustrate methodology in practice

How to read the guidance?

Each approach is presented in the guidance according to a fixed structure and contains the following:

- Description** Explaining in short terms the evaluation methodology in the context of ex-post evaluation of Effort Sharing-related policies.
- When to use it?** Describing the conditions or evaluation context under which the methodology can be appropriately applied (e.g. availability of evidence, stakeholders involved, evaluation questions).
- How to use it?** Describing the methodological steps to be taken in a concise and illustrative way.
- Advantages/disadvantages** Explaining & scoring the strengths and weaknesses of each methodology against a fixed set of characteristics. An explanation of the characteristics is given in the table below.

Characteristics	Score	Explanation
Data requirements	Low	How much data or evidence is required, to apply the method? Are many specific statistics, monitored data, ... required? ...
Complexity	Medium	How complex is the method? Are specific software requirements needed (other than Excel)?
Usefulness	High	How useful is the method to assess the evaluation criteria? E.g. decomposition analysis has low usefulness because can't directly be linked to one policy or measure.
Resources		How much time and other resources are needed to apply the method?
Evaluation criteria		What and how many of the evaluation criteria can be assessed by applying the method?
Communication / visualisation of results		Are the results easily communicated / visualised and thus raise understanding?

- Tools** What tools are publicly available online to support the evaluation methodology?
- Data sources** Which publicly available data sources or statistics, other than common sources (e.g. Eurostat), could support the evaluation?
- Related topics** Which other sections of the guidance relate to this topic? The other sections are indicated by means of a [HYPERLINK](#).
- Case studies of relevant examples of the approach** Which case studies or existing evaluation practices from Member States illustrate this evaluation methodology very well. The cases are indicated by means of a [HYPERLINK](#).

Want to know more?

- References to other, existing guidance documents or studies giving more explanation or interesting insights on the methodology.





Example on advantages/disadvantages

Decomposition analysis “can be used to quantify the influence of key factors on a variable of interest (e.g. CO₂ emissions). Such key factors can also be called drivers. As a consequence, decomposition analysis can help to understand why a variable of interest developed as it did, taking into account relevant drivers.”

Characteristics	Score	Explanation
Data requirements	Medium – High	Depending on complexity of governing function
Complexity	Medium	Excel will suffice, but setup and defining governing function may be complex
Usefulness	Low- Medium	Effects can only be indirectly attributed to specific policies and measures; For some sectors, circumstances may allow for medium usefulness. E.g. if only one policy has been introduced or policy has no overlapping effects, or results clearly indicate link to a specific policy (such as changing refuelling behaviour as in the example given above).
Resources	Low-Medium	The resources needed depend on the complexity of the governing function and whether various experts are needed for setting up the case under question. Software choice may also drive resource needs.
Evaluation criteria	Low	Effectiveness can only be indirectly attributed to specific policies and measures.
Communication / visualisation of results	High	Results can be visualised in an appealing and easily understandable manner. Decomposition methods that do not yield a residual are specifically useful for communication of results.

Agenda



Key characteristics of the guidance for ex-post evaluation



From a bibliography to a reviewed guidance



Design of the guidance? How to use the guidance?



Overview of methodologies



Where to find Member States' practices in the guidance? Case studies



Q & A



Methodologies for collecting evidence

Methodologies	Short description	Data requirements	Complexity	Usefulness	Resources	Evaluation criteria	Communication/ Visualisation
Surveys	Surveys are a common methodology for collecting large sets of data in a structured way from a sample. It is an effective way to gather new data, so typically used when there is a lack of existing primary data.	Low	Low	Medium	Low - Medium	High	High
Systematic literature review	A systematic review refers to a focused literature review that seeks to answer research question(s) using pre-defined eligibility criteria (inclusion/exclusion criteria) for documents and outlined and reproducible methods.	Low	Medium	Low - Medium	Medium	Low - High	Low - Medium
Focus groups	The focus group uses structured discussion that involves the progressive sharing and refinement of participants' views and ideas. It is well suited to cases where the views on evaluation topics are very divergent, but where discussion in groups may lead to a deeper viewpoint.	Low	Medium	Medium	Medium	High	Low
Interviews	Interviews are used to collect qualitative information in a (semi-) structured way and to collect the opinions of persons affected by a particular intervention. They have the most added value in an exploratory context, often as complement to a survey.	Low	Medium	Medium	Medium - High	High	Low
Monitoring performance data	Monitoring data are regularly collected, quantitative data about a policy and can include data relating to each component of the intervention logic model and to each evaluation criteria.	Low	Low-Medium	High	Medium - High	High	High



Methodologies for collecting evidence

Methodologies	Short description	Data requirements	Complexity	Usefulness	Resources	Evaluation criteria	Communication/ Visualisation
Surveys	Surveys are a common methodology for collecting large sets of data in a structured way from a sample. It is an effective way to gather new data, so typically used when there is a lack of existing primary data.	Low	Low	Medium	Low - Medium	High	High
Systematic literature review	A systematic review refers to a focused literature review that seeks to answer research question(s) using pre-defined eligibility criteria (inclusion/exclusion criteria) for documents and outlined and reproducible methods.	Low	Medium	Low - Medium	Medium	Low - High	Low - Medium
Focus groups	The focus group uses structured discussion that involves the progressive sharing and refinement of participants' views and ideas. It is well suited to cases where the views on evaluation topics are very divergent, but where discussion in groups may lead to a deeper viewpoint.	Low	Medium	Medium	Medium	High	Low
Interviews	Interviews are used to collect qualitative information in a (semi-) structured way and to collect the opinions of persons affected by a particular intervention. They have the most added value in an exploratory context, often as complement to a survey.	Low	Medium	Medium	Medium - High	High	Low
Monitoring performance data	Monitoring data are regularly collected, quantitative data about a policy and can include data relating to each component of the intervention logic model and to each evaluation criteria.	Low	Low-Medium	High	Medium - High	High	High



Analytical methodologies

Methodologies	Short description	Data requirements	Complexity	Usefulness	Resources	Evaluation criteria	Communication/ Visualisation
Indicator	Indicator analysis uses single or multiple indicators to track progress towards meeting an objective or obtaining a certain effect. It requires quantitative data that usually comes from a monitoring framework and represents a time series and/or different groups.	Medium – High	Low – Medium	Medium-High	Medium	High	High
Cost effectiveness	CEA will appraise a policy in terms of effectiveness in achieving single desired outcome for given level of cost, relative to its counterfactual. It is most useful when the objectives are clearly identified, connected to a quantified target and a clear baseline.	Medium	Low – Medium	Medium	Low – Medium	Medium	Medium
Cost benefit	CBA assesses whether a policy or measure is worth implementing (i.e. whether benefits outweigh costs) from a societal perspective. The major difficulty with CBA is to monetize all costs and benefits.	High	Low – Medium	High	High	Medium	Medium
Regression	This statistical method aims to investigate the relationship between two or more variables. It is a useful method to estimate impacts of a policy if (a) the policy effect can be characterised by a specific variable and; (b) good data is available on the trends for this variable as well as for the other variables.	High	High	Medium	Medium	Medium	Medium
Decomposition	Decomposition analysis can be used when one wants to quantify how various key drivers influence GHG emissions. These drivers do not directly depict policies and measures. The effects of policies will be indirectly visible through the changes in the drivers.	Medium – High	Medium	Low – Medium	Low – Medium	Low	High
Multi-criteria	In case where policy options may have different environmental and social impacts that are measured with different units. An MCA can provide a method for comparing different indicators and ranking the options while providing a transparent rationale for evaluation.	Low – Medium	Low	Low	Low	Medium	High



Analytical methodologies

Methodologies	Short description	Data requirements	Complexity	Usefulness	Resources	Evaluation criteria	Communication/ Visualisation
Indicator	Indicator analysis uses single or multiple indicators to track progress towards meeting an objective or obtaining a certain effect. It requires quantitative data that usually comes from a monitoring framework and represents a time series and/or different groups.	Medium – High	Low – Medium	Medium-High	Medium	High	High
Cost effectiveness	CEA will appraise a policy in terms of effectiveness in achieving single desired outcome for given level of cost, relative to its counterfactual. It is most useful when the objectives are clearly identified, connected to a quantified target and a clear baseline.	Medium	Low – Medium	Medium	Low – Medium	Medium	Medium
Cost benefit	CBA assesses whether a policy or measure is worth implementing (i.e. whether benefits outweigh costs) from a societal perspective. The major difficulty with CBA is to monetize all costs and benefits.	High	Low – Medium	High	High	Medium	Medium
Regression	This statistical method aims to investigate the relationship between two or more variables. It is a useful method to estimate impacts of a policy if (a) the policy effect can be characterised by a specific variable and; (b) good data is available on the trends for this variable as well as for the other variables.	High	High	Medium	Medium	Medium	Medium
Decomposition	Decomposition analysis how various key drivers influence GHG emissions. These drivers do not directly depict policies and measures. The effects of policies will be indirectly visible through the changes in the drivers.	Medium – High	Medium	Low – Medium	Low – Medium	Low	High
Multi-criteria	In case where policy options may have different environmental and social impacts that are measured with different units. An MCA can provide a method for comparing different indicators and ranking the options while providing a transparent rationale for evaluation.	Low – Medium	Low	Low	Low	Medium	High

Country case Luxembourg, decomposition analysis of road transport GHG emissions



Key evaluation issues and methodological challenges

Example on uncertainty analysis

While dealing with uncertainty for ex-post evaluations of policies *“two aspects are important: (1) quantifying the uncertainty of the evaluation results, and (2) reducing the uncertainty to improve the robustness of the evaluation findings. Both approaches are explained in the guidance.”*

How to quantify uncertainty? *“The influence of key variables on the result could be investigated by a sensitivity analysis. These key variables should be allowed to vary in order to test the robustness of the final result and should be linked to the drivers of the problem identified.”*



Example of partial sensitivity analysis

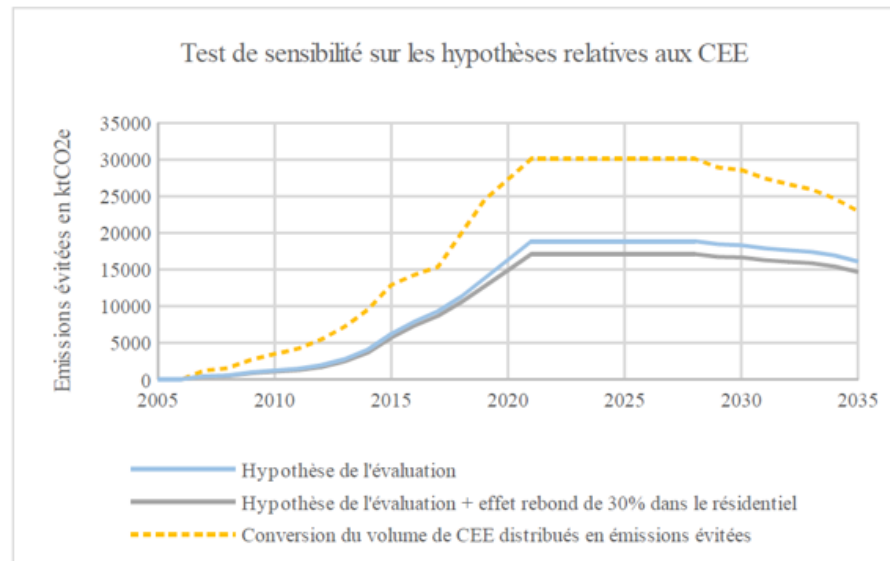
EVALUATION OF THE ENERGY EFFICIENCY CERTIFICATE SCHEME IN FRANCE (FRANCE, 2017)

When assessing the impact of the Energy Efficiency Certificate Scheme, it was recognised that there is a high level of uncertainty on the impact of the measures. This was related to the fact that detailed statistics were not available (uncertainty related to expected savings), the counterfactual scenario was unclear (especially the additionality of the measure compared to other measures) and indirect effects (most importantly the rebound effect) could not be quantified. Therefore a partial sensitivity analysis was done, performing the same calculation according to three different scenarios:

- One calculation corresponding with a best case scenario, assumed that all certificates resulted in avoided GHG emissions attributable to the measure;
- One calculation, the scenario selected as most likely, assumed that only 20% of energy and emission savings for the residential sector could be attributed to the Energy Efficiency Certificates, while the remaining part was attributed to other measures which provided a larger financial incentive. For other sectors, 100% of savings were allocated to the Energy Efficiency Certificates;
- One calculation based on the one above, but assuming a rebound effect of 30%.

This French case shows how a simple sensitivity analysis could be done and represented. The results give an insight into the uncertainty and how conservative the selected scenario is compared to alternatives.

Figure 17. Results of partial sensitivity analysis of French Energy Efficiency Certificate Scheme (France, 2017).





Where to find Member States' practices in the guidance? Case studies



Case #6: Economic evaluation of French *écopastille* (eco-tax bonus-malus & super-bonus new vehicles) 2008-2012 (CGDD, 2013)

Cost benefit analysis	Counterfactual analysis	Regression analysis
Rebound effects	Indicator analysis	Monitoring performance data

Description of policy measure

France set up in 2008 the bonus-malus and super bonus. The combination of these two schemes is an economic incentive for the acquisition and production of lower-emission vehicles. This has significantly changed the structure of passenger car sales in France, compared to what had been observed previously.

The **bonus-malus scheme** (BM) aims to stimulate the use of more fuel-efficient techniques in the automotive field by delivering a price signal that acts both on supply and on demand. It steers consumer choice towards more fuel-efficient vehicles, on the one hand; and it encourages manufacturers to make vehicles that meet this demand and innovate in this way, on the other hand. More specifically, the scheme must be able to accelerate the reduction of CO₂ emissions of new passenger cars by applying a tariff scale depending on CO₂ emissions. In 2008, a subsidy was granted to new cars emitting less than 120 gCO₂/km, while new vehicles emitting more than 160 gCO₂/km were taxed on the occasion of their first registration. The scale has gradually hardened over the years, for instance, in 2013 the subsidy is awarded below 105 gCO₂/km, whereas the penalty applies to vehicles emitting more than 135 gCO₂/km.

The **super bonus** aims to accelerate the rate of renewal of the vehicle fleet and thereby reduce its average emissions. In 2008, a premium of EUR 300 is granted, subject to the acquisition of a new vehicle emitting less than 130 gCO₂/km, for the disposal of a vehicle older than 15 years. As part of the recovery plan, the super bonus was replaced in 2009 and 2010 by a scrappage premium. Since 2011, the super bonus system is in place again. However, it was modified in 2012, where an amount of EUR 200 is granted, in addition to the ecological bonus when a vehicle older than 15 years is disposed.

Scope of the evaluation

The evaluation presents the main facts observed in France concerning the evolution of private new vehicle registrations between 2008 and 2012, some comparisons with previous years and with what was observed in other European countries. In the **cost benefit analysis** of the *écopastille* from 2008 to 2012, different pillars to examine its effects were touched upon: economic, environmental, and a socio-economic assessment.

Methodology

The evaluation performed a socio-economic cost benefit assessment, including the following **types of costs**:

- (1) Loss of consumer surplus: this is linked to a restriction of consumer's choice in comparison to their previous buying habits. In general, consumers do not value their costs over the total period of owning their vehicle, but only value these costs during the first years, namely from 3 to 5 years (in this French evaluation study, a period of 4 years is assumed).
- (2) The opportunity cost of public funds (coût d'opportunité des fonds publics, COFP): the deficit of the economic incentives and the shortfall of revenues for the government linked to the fall in the internal

consumption tax of energy products (TICPE – Taxe Intérieure de Consommation sur les Produits Énergétique, also formerly known as TIPP) associated with lower fuel consumption.

- (3) The effects in terms of local pollution: the *écopastille* scheme could favour diesel vehicles, which are more CO₂ efficient. However, local pollution emitted by a diesel vehicle is more harmful compared to a gasoline vehicle.

As for **the types of benefits**, the evaluation included:

- (4) Fuel savings: the report valued the reduction in fuel consumption over the lifetime of new passenger vehicles, removing the part of these savings (over 4 years) already "internalized" by consumers and therefore integrated into the 'loss of consumer surplus'.
- (5) The reduction of CO₂ emissions: a carbon value equal to 32 EUR/tCO₂ until 2010 was proposed, growing at a rate of 5.8% per year after that until 2030. This value makes it possible to monetise the gains of the bonus-malus scheme in terms of CO₂ emissions over the lifetime of the new vehicles.

So, in total 3 pillars are monetized: (i) economic, (ii) environmental, and (iii) socio economic. **Rebound effects** are also taken into account as a surplus of traffic is expected inducing social and environmental costs, which are not completely covered by the levies on road traffic (namely, tolls and TICPE).

In order to assess the impact of the eco-label scheme, a **reference scenario** was defined for the years 2008-2012 to isolate the effect of the scheme from the cyclical effects (counterfactual scenario). The assumptions to construct this reference situation are based on a comparison of the French situation with the EU on the one hand, and econometric studies on the other hand, for instance:

- Total CO₂ emissions are based on an econometric study (least squares regression) assessing the evolution of emissions as a function of the fuel price (price elasticity -0.35);
- Car registrations and the fleet are also estimated by econometric studies (least squares regression) based on the average price of private vehicles (price elasticities are -1.7 and -0.1513 respectively).

The table below presents the data sources and type of indicators applied for the cost benefits assessment.



On website: <https://effortsharing.ricardo-aea.com/>

Thank you! Any Questions?



Nele Renders

nele.renders@vito.be

Tom Dauwe

tom.dauwe@vito.be

VITO / EnergyVille

