

Climate change policy evaluations in the EU and Member States: Results from a meta-analysis

Authors

Noriko Fujiwara, Stefan Bößner
and Harro van Asselt

The authors are grateful to Douglas Fraser, Sebastian Voigt, Niki-Artemis Spyridaki, Michael ten Donkelaar, Emilie Alberola, Andreas Türk and Keith Williges for their contributions to the review of evaluations on a country basis.



The CARISMA Project started in February 2015 and received funding from the European Horizon 2020 programme of the EU under the Grant Agreement No. 642242. CARISMA intends, through effective stakeholder consultation and communication, to ensure a continuous coordination and assessment of climate change mitigation options and to benefit research and innovation efficiency, as well as international cooperation on research and innovation and technology transfer.

The sole responsibility for the content of this Discussion Paper lies with the authors. It does not necessarily reflect the opinion of the European Union. Neither the EASME nor the European Commission is responsible for any use that may be made of the information contained therein.

Abstract

This discussion paper presents the main findings from a meta-analysis of climate change mitigation policy evaluations in the European Union (EU) and six Member States: Austria, Czech Republic, France, Germany, Greece and the United Kingdom. In doing so, it seeks to provide insights into how evaluation practices might be improved, and responds to information and knowledge needs about the state of European climate change mitigation policies, which are expected to become ever more important in the context of the Paris Agreement and the forthcoming Governance Regulation. The study further aims to improve the understanding and to contribute to ongoing studies of European policy evaluation practices.

Our sample of 236 policy evaluations in the EU and six of its Member States covered the period 2010-2016. Evaluations tend to increase in the years before major policy events. Compared with the results of a previous meta-analysis carried out in 2008-2009, formal evaluations commissioned by government bodies have been on the rise in 2010-2016. Most evaluations focus on the effectiveness and goal achievement and usually forgo a deeper level of reflexivity and/or public participation in the evaluation process. The analysis also revealed the dominance of the energy sector in the sampled evaluations. The paper finds that while the EU and the six Member States have made some progress in reducing emissions and increasing the share of renewable energy sources in the energy mix, other sectors such as transport and buildings for energy efficiency lag behind those efforts. The low number or indeed the absence of any policy evaluations in the agriculture, waste or land-use sectors is an area for further investigation.

Table of contents

- 1 Introduction 3**
- 2 Climate policy evaluation in the European Union: an overview 4**
- 3 Methodology for the meta-analysis of evaluations 5**
- 4 Results of the meta-analysis of evaluations 8**
 - 4.1 General statistics 8**
 - 4.2 Timing of publications 9**
 - 4.3 Affiliation of the authors 10**
 - 4.4 Commissioning bodies 11**
 - 4.5 Sectors 12**
 - 4.6 Reflexivity 15**
 - 4.7 Evaluation methods 15**
- 5 Recommendations 18**
- References 19**



1 Introduction

The evaluation of climate change mitigation policy is crucial for understanding how well policies and measures work. Policy evaluation offers analysts insights in the functioning of policies and provides policy makers with much-needed information on how to improve them. In addition, policy evaluations can enhance the transparency of policy implementation, which is essential to gain citizens' support for those policies.

The importance – and necessity – of climate policy evaluation is underscored by the Paris Agreement adopted in December 2015. The Paris Agreement puts in place a process in which countries pledge, in five-year cycles, non-legally binding “Nationally Determined Contributions (NDCs)”. The main accountability mechanism to ensure countries live up to their promises consists of various review processes, notably a review of implementation by individual parties (i.e. through the ‘enhanced transparency framework’ of Article 13¹), and a review of global efforts towards long-term goals of the Agreement (i.e. the ‘global stocktake’ of Article 14).

Since the early 1990s, and in line with the requirements under the United Nations Framework Convention on Climate Change (UNFCCC), the European Union (EU) has gained significant experience in monitoring and reporting greenhouse gas emissions, as well as the policies put in place to reduce emissions. Over time, the Union has strengthened its internal transparency mechanisms by requesting Member States to report on the *impacts* of these policies. Member States now have to report their progress on climate change policies under the EU Monitoring Mechanism Regulation (525/2013) (European Union 2013), due to be replaced by the Regulation on the Governance of the Energy Union, while other policy areas such as renewable energy and energy efficiency also require Member States to submit national action plans and reports (Umpfenbach 2015; Schoenefeld et al. 2016).

Ex-post climate policy evaluations can offer insight into a variety of aspects, including the amount of greenhouse gas emission reductions, the cost-effectiveness of policies, their social acceptance, or the coherence with other policies. However, systematic analyses of the information made available by such evaluations are still largely lacking, with the notable exception of Huitema et al. (2011), which reports on a meta-analysis of 259 evaluations, covering the period from 1998-2007 for the EU and several EU Member States.² This discussion paper offers a recent application of this type of analysis, focusing on *ex-post* climate policy evaluations, to reflect more recent developments and gain updated insights into climate policies.

¹ Parties to the Paris Agreement are to report not only through the enhanced transparency framework on their emissions trends (through annual greenhouse gas inventory) but also through biennial reports that need to indicate how much progress has been made in implementing and achieving their nationally determined contributions under the Agreement (see van Asselt et al. 2017).

² The study by Huitema et al. (2011) was carried out by universities and research institutes under the EU Seventh Framework Programme project: Adaptation and Mitigation Strategies for Europe (ADAM).

Through careful and systematic analysis of policy evaluations, the CARISMA project aimed to enhance understanding of existing climate policy evaluation practices in the EU and Member States.³ This paper first provides an overview of the EU evaluation practices, clarifies the methodology, then summarises the results of a meta-analysis for discussions and suggests recommendations for EU policymakers.

2 Climate policy evaluation in the European Union: an overview

Since 2002, the European Commission has been committed to the EU agenda of Better Regulation,⁴ which highlights, among others, the *ex-ante* impact assessments of policy initiatives, the monitoring and *ex-post* evaluations of the existing policies as well as the importance of stakeholder consultation in these processes.⁵ More recently, in 2012, the European Parliament introduced *ex-ante* impact assessments by establishing a dedicated service within its administration. Since 2013, *ex-post* evaluations have been added to complete an entire legislative cycle from agenda setting to scrutiny of legislative proposals (Anglmayer 2016). The growing recognition within the EU of the value of the evaluation process has resulted in an increasing demand for the evaluation of environmental policies and programmes (Mickwitz 2013), including evaluations in the area of climate policy.

Like policy evaluation in general (Nilsson et al. 2008), it has been challenging to evaluate climate policies as it can be hard to identify clear outcomes of policies, and policies often interact with other policies (Mickwitz and Birnbaum 2009; Haug et al. 2010; Mickwitz 2013). Within the EU, these barriers to an effective evaluation process are compounded by the complexity of the governance system (Mickwitz 2013). Moreover, there are important political barriers for further strengthening evaluation practice in EU Member States. For instance, they require financial resources that governments may be unwilling to allocate, and Member States may be unwilling to cede more powers to EU institutions for this function (Hildén et al. 2014).

One of the most relevant pieces of EU legislation for climate change mitigation policy evaluations is the Monitoring Mechanism Regulation (MMR) (European Union (EU) 2013).⁶ The MMR requires Member States to report “quantitative estimates of the effect of policies and measures on emissions by sources and removals by sinks of greenhouse gases” (Article 3.2(a)(v), EU 2013) and to report the following elements in their information on policies and measures (Article 13.1(c) (iii)-(vii), EU 2013): the status of implementation of the policy or measure or group of measures; indicators to monitor and evaluate progress over

³ This paper explains the scope of the meta-analysis. The actual performance of climate change policies based on the policy evaluations is discussed in Fujiwara et al. (2017).

⁴ <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:I10108>

⁵ https://ec.europa.eu/info/law/law-making-process/better-regulation-why-and-how_en; see also EEA (2016a); and Ashford and Renda (2016:31).

⁶ The Regulation 525/2013/EU (European Union 2013) is an update of the EU’s Monitoring Mechanism Decision (Decision 280/2004/EC; European Union 2004). The MMR is not the only EU legislation that calls for the evaluation of the effects of policies and measures. As Hildén et al. (2014) note, other Directives – e.g. Article 22.1 of the Renewable Energy Directive (2009/28/EC) and Article 24.1 of the Energy Efficiency Directive (2012/27/EU) – likewise call for Member States to report on progress made in the implementation.

time; quantitative estimates (both *ex-post* and *ex-ante* assessments) of the effects of policies and measures on emissions by sources and removals by sinks of greenhouse gases; estimates of the projected costs and benefits of policies and measures, as well as estimates of the realised costs and benefits of policies and measures; and all references to the assessments and the underpinning technical reports. While these provisions are encouraging in that they call on Member States to provide both *ex-ante* and *ex-post* information on the effects of mitigation policies, and also encourage Member States to offer estimates of *ex-ante* and *ex-post* costs and benefits, they also leave much discretion to the Member States, as indicated by the various mentions of the words “where appropriate” and “where available” (Article 13.1(c) (v)-(vii), EU 2013), as well as by the fact that Member States can opt to assess the effects of a group of measures. Initial reviews of reporting practices suggest that Member States thus far have hardly included *ex-post* assessments of the effects of policies and measures in their reports (Hildén et al. 2014; Schoenefeld et al. 2016).

Indeed, the necessary capacity to carry out *ex-post* evaluations is not fully developed throughout the EU. A 2009 study carried out in preparation for the MMR found that the EU15 Member States tended to have more experience in *ex-post* evaluations and more often have formalised monitoring and evaluation systems in place than the newer Member States (AEA et al. 2009: 14). Even though the situation might have changed since then, the study suggests that not all the EU Member States have the full capacity to carry out *ex-post* evaluations.

All this is not to say that *ex-post* evaluations are not available for the EU and its Member States. Indeed, the European Environment Agency (EEA) seeks to go beyond formal evaluation procedures such as those carried out by (or on behalf of) the European Commission in the context of the Better Regulation agenda, adding value by evaluating policies within a more environment-specific context as well as those policies influencing environmental policies⁷ according to its autonomous mandate. It has also developed a conceptual framework for policy evaluation that builds on key policy evaluation criteria (EEA 2016).

3 Methodology for the meta-analysis of evaluations

The CARISMA project team carried out a meta-analysis of climate policy evaluations. A meta-analysis (or ‘systematic review’) allows for the synthesis of a large body of information (Vedung 2010), in this case climate policy evaluations.

As a starting point, the project team made several important choices concerning the scope of the meta-analysis.

⁷ The EEA seeks to hold a dialogue about policies on changes in ecosystems, the production and consumption system, or the food, energy and mobility systems and engage in such a dialogue with the EEA member countries and the European Environment and Information Network (Eionet), European institutions, the environment evaluators community and interested evaluation professionals (EEA 2016: 4).

The first choice was *which geographical jurisdictions* to include. Examining 28 Member States would have been challenging, considering time and resource constraints as well as language barriers. A previous meta-analysis of climate policy evaluations focused on the EU as well as six Member States: United Kingdom, Germany, Italy, Finland, Portugal, and Poland (Huiteima et al. 2011). Following this model and drawing on the local expertise of CARISMA project partners, the project team opted to include the three largest EU Member States (France, Germany, and the UK) as well as smaller Member States from Central and South-Eastern Europe (Austria, Czech Republic, and Greece). In addition, since important climate policy evaluations for the EU had been carried out at the EU level, the EU was included as a separate jurisdiction.

The second scope-related choice concerned the *time period of the analysis*. The meta-analysis by Huiteima et al. (2011) covered evaluations from January 1998 to March 2007. To offer a more recent update, reflecting climate policy developments in the period after the COP15 in Copenhagen and after the enactment of the EU's 2020 climate and energy package, both in 2009, the CARISMA project team decided to cover the period from January 2010 to December 2016.

The third choice concerns the *eligibility of evaluations for the analysis*. The number of climate policy evaluations is potentially large, and the project team sought to limit the number of evaluations in several ways:

- The team only examined *ex-post* evaluations, including studies that have both *ex-ante* and *ex-post* elements, but excluded purely *ex-ante* evaluations.
- The team only examined climate change mitigation, not adaptation, policies.
- The team only examined evaluations of policies reported as climate policies by cross-checking with reports submitted to the UNFCCC (e.g. the latest National Communications under the Convention). However, it should be noted that it might not always be clear whether a certain policy is considered a climate change mitigation policy or not. Some policies reported by Parties may be of a general nature (e.g. a country may have a general forest policy, which may in part promote sustainable forest management, leading to greenhouse gas emission removals) or may only indirectly contribute to climate change mitigation (e.g. urban planning policies that may encourage modal shifts). Evaluations were however included, when they included a reference to specific mitigation policies.
- The team excluded purely academic articles. Although this might exclude genuine evaluations that met all other criteria, it also avoids overlap (as some articles may be based on working papers or reports that the project team covered).
- The team excluded non-systematic analyses. This meant, for instance, that position papers by NGOs, industry groups, or trade associations were excluded. While such papers may include evaluations of policies (emphasising aspects that might escape

scrutiny in 'formal' policy evaluations⁸), their primary purpose is not considered to be evaluation as such, but rather advocacy for policy change.

- The team only looked at EU-level and national (as opposed to also subnational) climate policies.
- The team only looked at published documents that are made available to the public.

Of course, these choices can influence the results of collecting and aggregating evaluations. For instance, excluding academic articles not only means that the number of evaluations covered is lower than it otherwise could have been; but it may also influence the number of 'independent' evaluations.

The next step was for the project partners to carry out the meta-analysis to collect all relevant evaluations for a specific jurisdiction that met the proposed selection criteria. Specifically, the partners would: (1) check the websites of national governments; (2) examine climate policy reports submitted to the EU and UNFCCC; (3) carry out searches in Google Scholar; (4) visit websites of key national and international environmental consultancies; and (5) request further information from relevant stakeholders (including other researchers). For those countries where only few evaluations could be found, the project partners would carry out further searches. Each project partner would include basic information (e.g. authors, title, year, sector) on the evaluations they collected in an Excel sheet. This formed the total sample of the meta-analysis.

The third step was to finalise the eligibility criteria and collect key information from (i.e. code) the evaluations, to create a more comprehensive set of information on the basis of a common template (see Appendix 1 in Fujiwara et al. 2017). The template was developed drawing on the template used by Huitema et al. (2011). It was updated and adjusted following internal feedback from CARISMA project partners.

The collection of information for each evaluation facilitated the following step, which was to aggregate information collected in the templates. Information collected for each of the evaluations was aggregated, focusing on: (1) the year of publication; (2) the affiliation of authors; (3) commissioning bodies of evaluations; (4) sectoral coverage; (5) the nature of the evaluation (reflexivity); (6) evaluation methods used; (7) evaluation criteria used; and (8) whether political recommendations were made. This provided the basis for results reported in the next section. The aggregation process also helped to verify information collected by the project partners, spot gaps and inconsistencies, and, if needed, ask for further clarification from the partners or exclude the evaluations that, on closer inspection, did not meet the agreed eligibility criteria.

While the meta-analysis allows the project team to derive findings about the climate policy evaluation practice, a number of caveats should be made. First, the number of evaluations

⁸ Policy evaluation theory distinguishes between 'formal' and 'informal' evaluations, based on the actors involved in the evaluation (Schoenefeld and Jordan 2017, drawing on Weiss 1993). The key distinction here is that formal evaluations are carried out or driven (e.g. commissioned) by governments, or those responsible for the policy, whereas informal evaluations are driven by other societal actors.

found is likely to be non-exhaustive, even though the project team took steps to ensure a wide coverage of evaluations. It is likely that at least some evaluations that could be considered relevant were not covered. This may be for the reasons sketched above (e.g. because evaluations were not publicly available), but it may also be because it was unclear whether an evaluation addresses a mitigation policy. Second, the answer to some of the questions asked to the project partners is inevitably subjective. For example, evaluations do not always clearly spell out which criteria or methods are used. Moreover, whether an evaluation is truly “reflexive” is not always clear. The project team sought to address this to the greatest extent by offering as much guidance as possible to the coders (see Appendix 1 in Fujiwara et al. 2017). Like the previous meta-analysis (Huitema et al. 2011), this task involved multiple project partners working in a decentralised manner, and some discretion was left to the individual coders on practical choices. Consequently, despite the coordination efforts from the early stages of work, it was not possible to ensure full consistency in the ways coding was done in each jurisdiction. Furthermore, the large variety of methods and criteria used for assessing policies as well as the different evaluation practices in different Member States make a comparative substantive analysis difficult.

4 Results of the meta-analysis of evaluations

The following sections report and analyse results of coded evaluations in a way comparable to Huitema et al. (2011), applying the above-mentioned methodology. This approach enables us to compare new results of evaluations published in 2010-2016 with those in 1998-2007 (Huitema et al. 2011).

4.1 General statistics

The total number of sampled evaluations is 236. The three major jurisdictions contributing to the evaluations are the EU (70 evaluations, 30%), the UK (64 evaluations, 27%) and Germany (59 evaluations, 25%). Four countries, Greece (20 evaluations), Austria (8 evaluations), France (10 evaluations) and the Czech Republic (5 evaluations) together account for the remaining 18%.⁹ This leaves the question why the number of evaluations is relatively low in France. The variation in number of evaluations implies a discrepancy in evaluation practices across Member States for reasons other than limited capacity as discussed earlier. Lists of sampled evaluations can be found in Appendix 2 of Fujiwara et al. (2017). Table 1 gives an overview.

⁹ The Member State shares of total greenhouse gas emissions in the EU28 in 1995-2015 were the following: Germany 20.8%, UK 12.1%, France 10.7%, Czech Republic 2.9%, Greece 2.2%, and Austria 1.8%. http://ec.europa.eu/eurostat/statistics-explained/index.php/Greenhouse_gas_emission_statistics.

Table 1. Evaluations per jurisdiction

	EU total	DE	UK	GRE	AT	FR	CZ	Total
Total number of evaluations	70	59	64	20	8	10	5	236

Source: Authors' compilation of data.

In comparison, the sample size of evaluations covered by Huitema et al. (2011) – which also included adaptation policies and academic articles – was 259, ranging from the EU (105 evaluations) and the UK (78 evaluations) to Portugal (10 evaluations) and Poland (6 evaluations). The diversity of the new sample for 2010-2016 is similar to those of the old sample for 1998-2007, although the average number of evaluations (compared to the number of years covered) is comparatively higher (33.7 evaluations/year compared to 25.9 evaluations/year).

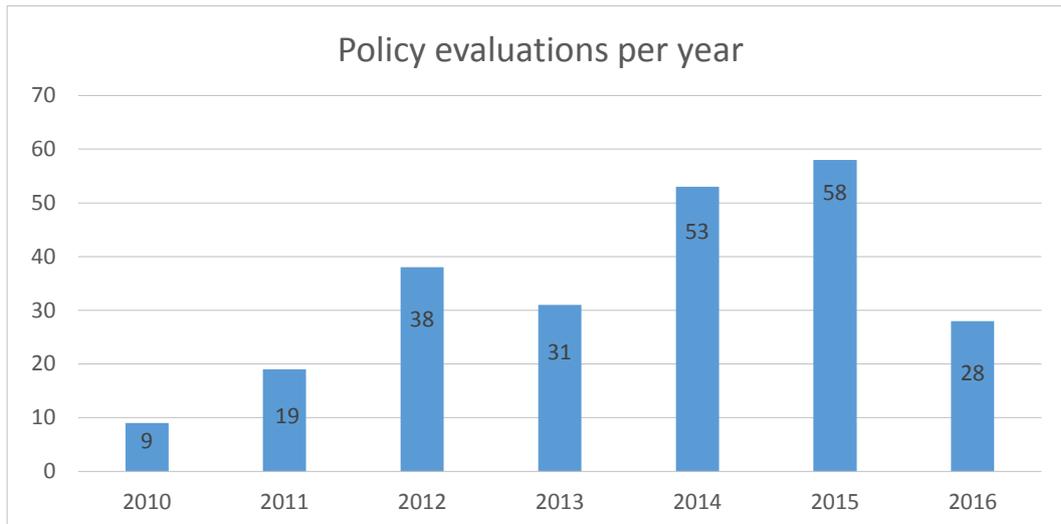
4.2 Timing of publications

The number of evaluations continued to increase towards 2015, except for 2013, and then declined by more than half (Figure 1). It is unclear whether the year 2016 is exceptional or signals a changing trend. It is possible that a number of evaluations were completed in 2016 but not yet published.

Although the project team did not systematically investigate a causal relationship between the timing of publications and the number of evaluations, the number of evaluations could be said to be linked to specific policy developments and requirements at international, EU and national levels. The evaluations peaked in 2014-2015 after a significant increase in 2012. Nearly half (47%) of the total evaluations were published in the years 2014 and 2015. The two years, 2014 and 2015, can be regarded as important milestones to evaluate the effects of existing policies in preparation for two major policy events. One is the submission of intended nationally determined contributions (INDCs) to the UNFCCC, which were due in the first quarter of 2015.¹⁰ The other is the launch of the Energy Union in Europe (European Council 2015). The evaluations in the third highest year, 2012, include the first carbon market report at the end of the second phase of the EU ETS, and sectoral analyses of non-ETS sectors.

¹⁰ http://unfccc.int/focus/indc_portal/items/8766.php.

Figure 1: The number of evaluations per year



Source: Authors' compilation of data.

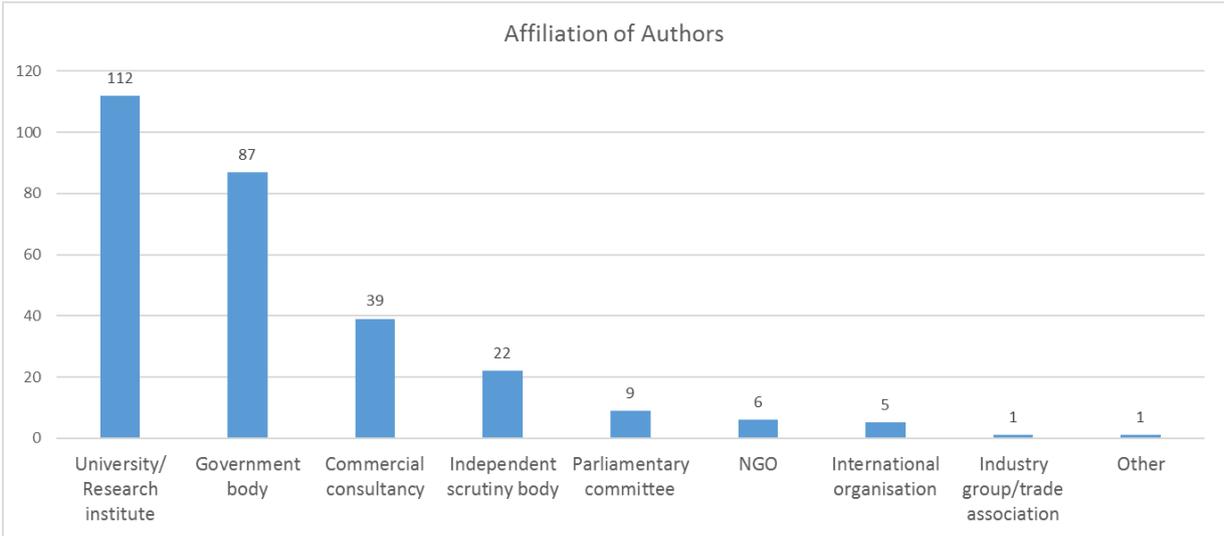
The new sample for 2010-2016 remained mostly in the range of around 30-60 evaluations per year. This falls between the low level of evaluation activities (0-30 evaluations per year) in 1998-2004 and the high level of activities (80-90 evaluations per year) in 2005-2006 (Fig.1, Huitema et al. 2011).

4.3 Affiliation of the authors

While the largest group of affiliated authors was *universities and research institutes*, *government bodies* were the second largest group of affiliated authors. Figure 2 shows the affiliation of authors, highlighting that *universities or research institutes*, followed by *government bodies*, were responsible for the clear majority of evaluations.¹¹ It is possible, however, that other groups, especially NGOs or trade associations, were underrepresented in the sample because our eligibility criteria excluded position papers. *Government bodies* rather than *universities or research institutes* only contributed most to evaluations as authors in the EU and the Czech Republic.

¹¹ Evaluations may be counted under more than one category.

Figure 2. Affiliation of the authors



Source: Authors’ compilation of data.

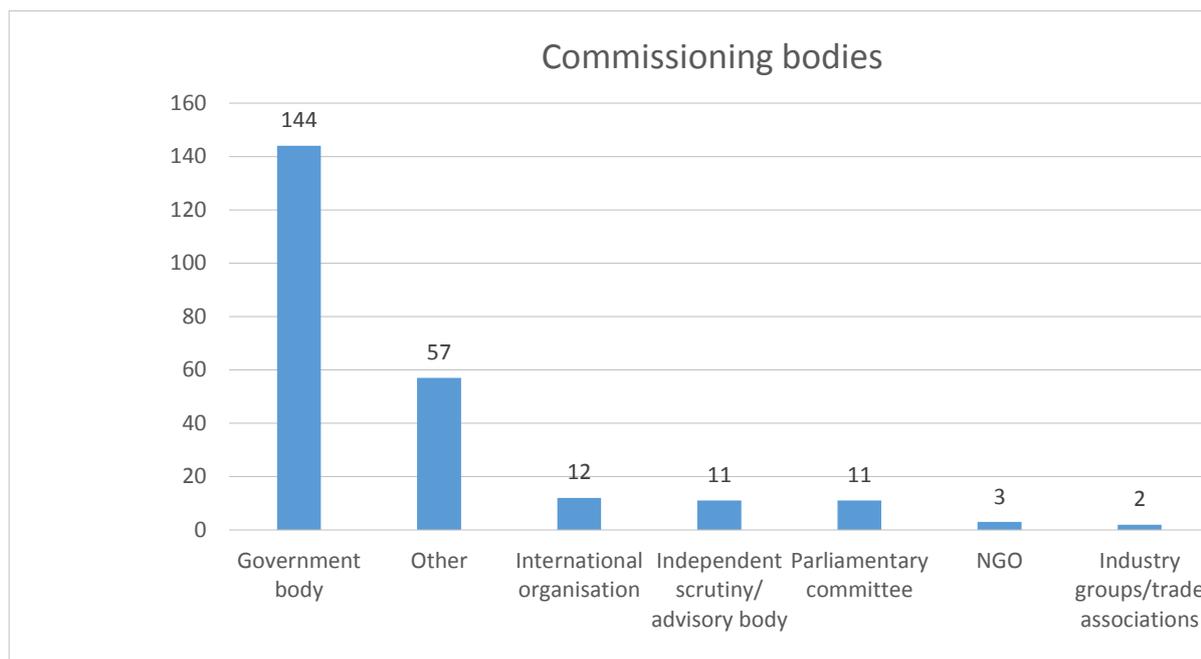
The previous study by Huitema et al. (2011) also ranked universities or research institutes (about 135 evaluations) on top, but followed by commercial consultancies (50-60 evaluations), and international organisations (20-30 evaluations) (Figure 2 in Huitema et al. 2011). A fewer number of government bodies authored policy evaluations.

4.4 Commissioning bodies

Figure 3 shows that most evaluations were commissioned by *government bodies*.¹² The second highest number of evaluations does not fall under any of the specified categories, i.e. *others*. This category may include any evaluations for which the commissioning bodies could not be identified. In contrast to Member State jurisdictions, the majority of EU evaluations were commissioned by *government bodies*. It is possible that NGOs were underrepresented in the sample due to the eligibility criteria. For example, certain *ex-post* evaluations carried out for their internal purposes may well have been excluded from the sample.

¹² Evaluations may be counted under more than one category.

Figure 3. Commissioning bodies



Source: Authors' compilation of data.

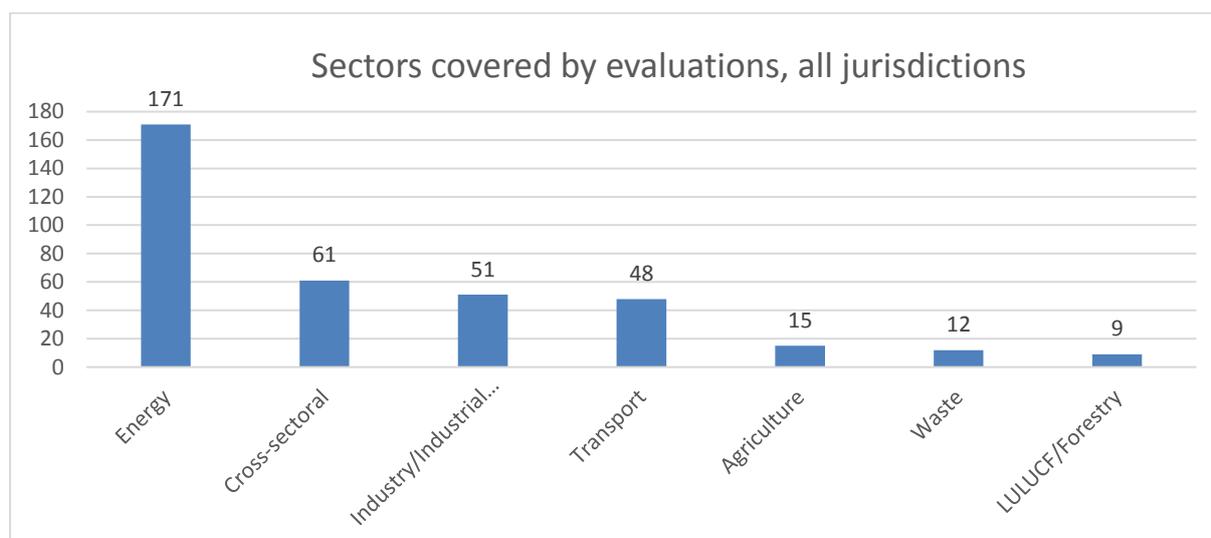
Huitema et al (2011) do not provide details about the breakdown of commissioning bodies but differentiates whether the relevant evaluation was commissioned or not. Thus the new sample for 2000-2016 cannot be adequately compared with the study with respect to this question.

4.5 Sectors

Figure 4 reveals that evaluations in the energy sector are dominant in the sample,¹³ with 171 evaluation entries covering the sector (multiple answers were possible). This was followed by cross-sectoral evaluations (61 entries), industry/industrial processes (51) and transport (48). This pattern is common to most of the jurisdictions covered, except the EU, which had a relatively higher share of cross-sectoral approaches than Member States. Moreover, some sectors, such as the agricultural, waste and land use, land-use change and forestry (LULUCF) sectors, are clearly under-represented. The LULUCF evaluations have been particularly scarce, possibly because the sector was not accounted for in the EU-wide emission reduction target up to 2020.

¹³ Evaluations could be – and have been – counted under more than one category, which is the reason for the divergence of the total number of evaluations in our sample (236) and the total number of evaluations in Figure 4 (367).

Figure 4. Sectors covered, all jurisdictions



Source: Authors' compilation of data.

Looking at specific jurisdictions, the share of the energy sector is particularly high in the UK and Germany (Figure 5).¹⁴ The sectoral spread in the sampled evaluations can be compared with the sectoral shares of actual greenhouse gas emissions in the EU28 in 2015, i.e. energy (55%), transport (23%), industry (8%), agriculture (10%), and waste (3%).¹⁵ The aggregated data suggests that both the energy and industry sectors are more represented than their actual shares, despite the challenges of delineating specific sectors,¹⁶ and the multiple counting. By contrast, transport, agriculture and the LULUCF sectors are under-represented given their actual shares of emissions.

The strong focus on the energy sector in climate change mitigation policies corresponds to observations made in the CARISMA analysis of information available in international climate change mitigation policy databases (Böbner et al. 2017). Of all the databases analysed, the large majority contained information about energy policies but only a fraction conveyed any information on mitigation policies in the agricultural or waste sector (Böbner et al. 2017).

Possible explanations could include the following:

On the energy sector:

¹⁴ Evaluations may be counted under more than one category. Also, UK waste and agricultural policies have been addressed by a few cross-sectoral studies and were classified as such.

¹⁵ [http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Greenhouse_gas_emissions,_analysis_by_source_sector,_EU-28,_1990_and_2015_\(percentage_of_total\)_new.png](http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Greenhouse_gas_emissions,_analysis_by_source_sector,_EU-28,_1990_and_2015_(percentage_of_total)_new.png).

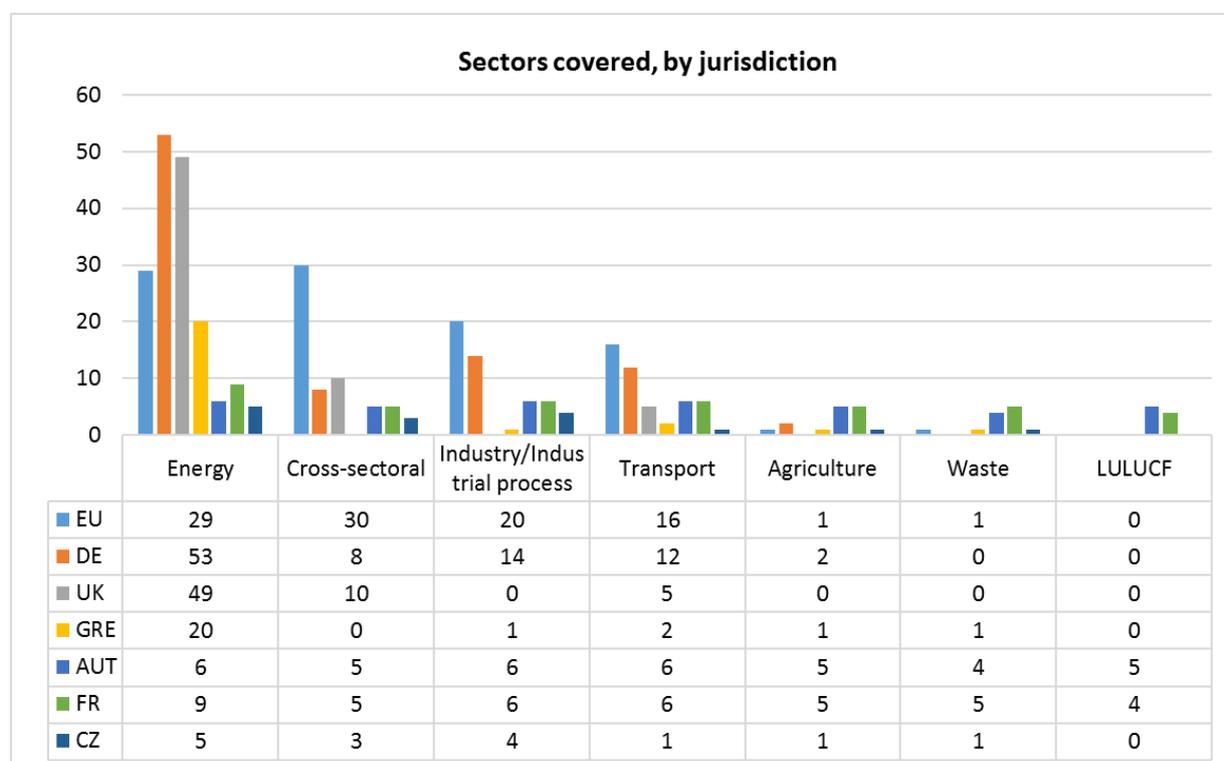
¹⁶ However, grouping emissions into different sectors is not standardised; the European Commission places all "combustion from fuels" under the energy category including "combustion from construction and manufacturing" which might be placed under the "industrial processes" category by other institutions.

- The energy sector is the sector responsible for the largest share of emissions in Europe, and has the largest mitigation potential;¹⁷
- The sector can be easily measured, monitored, quantified and verified;¹⁸;
- It is considered to deliver emission reductions in a more cost-effective way than in other sectors such as transport (e.g. “to promote reductions of greenhouse gas emissions in a *cost-effective and economically efficient* manner” EU 2003; see also EU 2009a); and
- It has additional mitigation potential through energy savings by end-users (EU 2012; EU 2009b).

On the industry sector:

- As far as the over-representation of the industry sector is concerned, one explanation might be the sensitivity of the sector for the overall economy.

Figure 5. Sectors covered, by jurisdiction



Source: Authors' compilation of data.

Huitema et al. (2011) do not provide for a sectoral breakdown per country, meaning that a direct comparison is not possible. Nevertheless, it is noteworthy that in the sample, evaluations carried out on the EU level seem to roughly reflect the cross-sectoral

¹⁷ The energy sector accounted for 55% of the total greenhouse gas emissions in EU28 in 2015. [http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Greenhouse_gas_emissions,_analysis_by_source_sector,_EU-28,_1990_and_2015_\(percentage_of_total\)_new.png](http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Greenhouse_gas_emissions,_analysis_by_source_sector,_EU-28,_1990_and_2015_(percentage_of_total)_new.png).

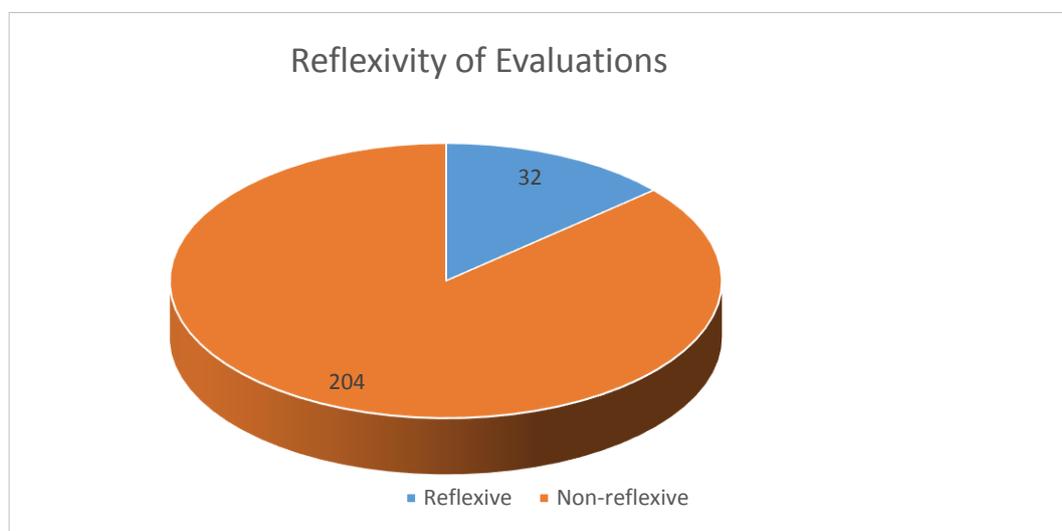
¹⁸ The energy sector has been covered by the EU emissions trading system (ETS) since 2005. See EU (2003) and EU (2009a).

distribution of actual emissions,¹⁹ while all Greek evaluations focus primarily on the energy sector.

4.6 Reflexivity

Policy evaluations can be conducted in a reflexive or non-reflexive manner. While the latter approach entails answering the question whether objectives of a given policy were reached according to certain criteria, a reflexive policy evaluation questions the objective and the chosen means to reach this objective critically and tries to address questions like whether the policy itself was/is justified. In the sample, the majority of the evaluations (204 evaluations, 86%) are found to be *non-reflexive* (Figure 6).

Figure 6. Reflexivity of evaluations by all jurisdictions



Source: Authors' compilation of data.

The high share of non-reflexive evaluations is comparable to the share (82%) in the 1998-2007 sample (Huitema et al. 2011).

However, a closer look at each country shows mixed results across jurisdictions. On the one hand, the EU, Germany, the UK and Austria have the highest shares of *non-reflexive* evaluations at 100% (70 evaluations for EU and 59 evaluations for Germany), 83% (53 out of 64 evaluations for UK) and 75% (6 out of 8 evaluations for Austria). On the other hand, Greece has an even split (10 evaluations each) while France and the Czech Republic have *reflexive* evaluations at 60% (6 out of 10 evaluations for France; 3 out of 5 evaluations for the Czech Republic).

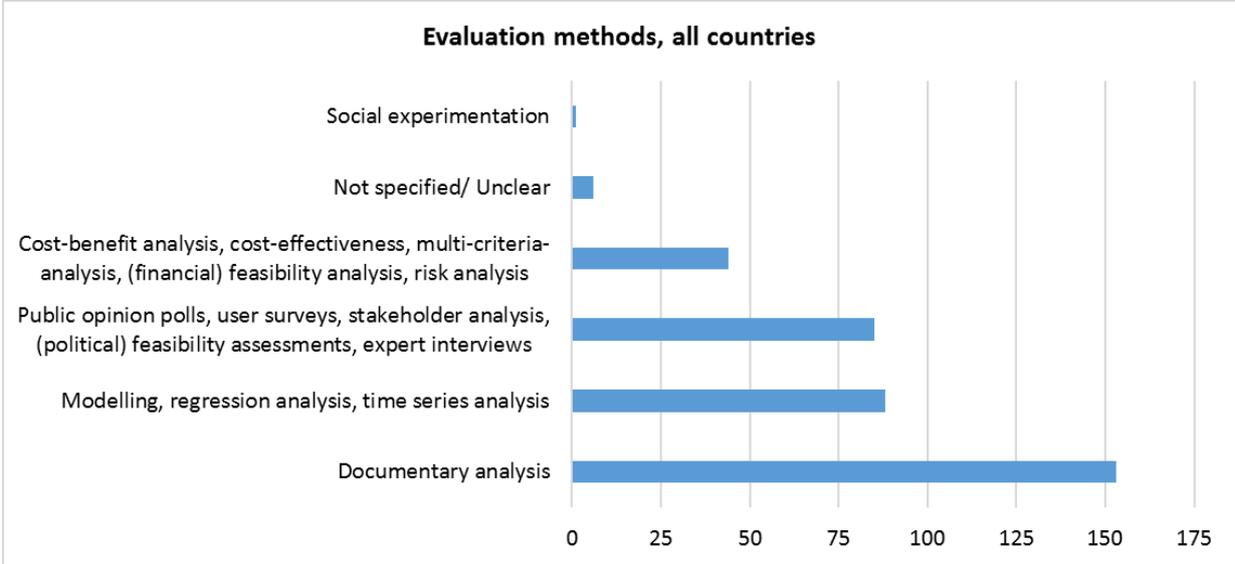
4.7 Evaluation methods

Using the evaluation methods and evaluation criteria as a proxy for reflexivity, the majority of evaluations were found to be non-reflexive. The largest number of evaluation methods applied (multiple answers were allowed) belonged to the category of *documentary analysis*

¹⁹ Evaluations in France and Austria also seem quite evenly distributed, but the sample size was quite small.

(153 entries), followed by *modelling, regression analysis or time series analysis* (88), *public opinion polls, user surveys, stakeholder analysis, feasibility assessments or expert interviews* (85) and *cost benefit analysis, cost-effectiveness, multi-criteria analysis, feasibility analysis or risk analysis* (44) (Figure 7).²⁰ In terms of stakeholder involvement, it can be assumed that most of the methods applied are neither participatory nor interactive, except for the second category.

Figure 7. Evaluation methods, all jurisdictions



First, the number of evaluations is not stable, but seems to be linked to specific climate policy developments and events at the national, European, and international levels. For example, evaluations increased significantly before countries first published their INDCs, and prior to, and in the wake of, the European Commission’s launch of its Energy Union initiative. This implies that international climate policy events – including not only the regular and review reporting already part of the UNFCCC regime, but also the new global stocktake starting in 2023 and its predecessor, the Talanoa Dialogue due in 2018 – may influence climate policy evaluation activities by setting milestones. Moreover, the five-yearly preparation of NDCs by all parties to the Paris Agreement is likely to spur climate policy evaluations, as evaluations can help the EU as a whole and Member States determine what level of ambition is adequate for their future policies.

Second, the largest group carrying out the sampled evaluations were universities and research institutions. However, a relatively large number of evaluations were carried out by government bodies, showing a significant increase from the sample of Huitema et al. (2011). Our analysis also found that government bodies are responsible for commissioning the large majority of evaluations in the sample. This suggests that while a large part of evaluations continues to be of an informal nature, formal evaluations may be on the rise.

Third, the dominance of the energy sector in evaluations is striking but not surprising, given the mitigation potential of the sector and the availability of methods to evaluate

²⁰ Evaluations may be counted under more than one category.

mitigation policies in the sector. However, the low number or indeed the absence of any policy evaluations in the agriculture, waste or land-use sectors may be an area of further investigation. Emissions from these sectors are still covered less by EU mitigation policies because until recently the EU emission reduction targets did not take into consideration the LULUCF sector, which means the sector was outside the scope of the major policy initiatives for the period until 2020.²¹ Or the nature of the sectors may mean that mitigation policies (and policy evaluations) are subsumed under larger policy initiatives (e.g. on sustainable agriculture or sustainable forest management) where climate change mitigation is but one of several policy goals. In any case, in the EU where, for example, the agricultural sector represents about 10% of greenhouse gas emissions,²² it is important to understand how well policies covering different aspects of these emissions have worked individually and how different policies have influenced each other, e.g. how agriculture, rural development, energy, and climate policies affect each other on bioenergy or biofuel production.

Fourth, the sample revealed that the vast majority (more than two-thirds) of the evaluations are not reflexive or participatory, confirming the findings in Huitema et al. (2011). Besides the small number of reflexive evaluations (i.e. those evaluations which examine the policy and its objectives more critically), most evaluations assessed policies against the criteria *effectiveness and goal attainment* and/or *cost-effectiveness*. Section 4 further suggests that, with the exception of the UK and Greece, evaluations hardly addressed questions related to the fairness or legitimacy of policies. While methods such as 'public opinion polls' and 'stakeholder analyses' have been well integrated in the EU evaluation practices and carried out relatively often, simple documentary analysis as well as modelling efforts still remain the methodology of choice for most (45%) evaluations assessed. This suggests that a large part of evaluations are not quite participatory in nature. One possible explanation for this finding may be the exclusion of position papers by NGO and lobby groups.

However, using evaluation criteria other than effectiveness/goal attainment or cost-effectiveness and participatory methods in evaluation is fraught with difficulties. For instance, assessing the fairness requires, first of all, the establishment of a benchmark of what can be considered 'fair' and – if the evaluation is to allow for comparisons – such benchmarks would need to be consistently applied. By contrast, the benchmarks for evaluations of effectiveness (e.g. tons of CO₂ emissions reduced) or cost-effectiveness (e.g. costs/ton of CO₂ emissions reduced) can – though need not²³ – be relatively straightforward. In other words, the application of some criteria may involve important (subjective) choices on the part of the evaluator,²⁴ which may make it more difficult to allow for comparative analyses. The use of participatory methods also faces particular

²¹ https://ec.europa.eu/clima/policies/forests_en.

²² http://ec.europa.eu/eurostat/statistics-explained/index.php/Agriculture_-_greenhouse_gas_emission_statistics#Agriculture.27s_contribution.

²³ Although policy goals may seem undisputed, the policy goal as formulated may mask underlying contestations between different societal actors about what a specific policy should achieve.

²⁴ While this finding may apply to all criteria to some extent, it applies more strongly to those criteria where metrics for evaluations (e.g. amount of CO₂ emissions reduced; costs per unit of CO₂ emission reduction) are absent.

challenges, including the costs of stakeholder engagement, and the needs for avoiding bias and ensuring representativeness (i.e. who participates).

Lastly, less than half of the evaluations presented political recommendations. There is a significant variance in the share of such recommendations between Member States on the one hand, and the EU on the other. The presence of political recommendations may depend on the role of evaluations envisaged in the relevant jurisdiction, specifically how far policy evaluations should go beyond the technical level, and how such evaluations should contribute to the legislative processes (e.g. providing evidence to policymakers in a closed process or directly submitting them to an open legislative process).

These findings have to be interpreted with care, due to limitations related to the eligibility criteria (see Section 3), such as the exclusion of academic publications, documents that are not publicly available and subnational policy evaluations. However, this new meta-analysis and the previous one enables researchers to track the long-term trend in almost 20 years and understand the diversity of policy development across jurisdictions and sectors. This study does show that while evaluation of mitigation policies is quite advanced in some jurisdictions, there is still some room for improvement, not only in terms of the quantity of the evaluations but also in terms of their quality.

5 Recommendations

What insights does this meta-analysis provide for policymakers? First, the meta-analysis reported here and in Huitema et al. (2011) show that a wealth of climate policy evaluations is being produced on an annual basis as a subset of policy evaluations in Europe. While the evaluation practice has not been compared across case studies (e.g. how does Germany evaluate mitigation policies compared to the United Kingdom), the large and increasing number of evaluations might harbour some redundancies and overlaps. The exercise of identifying, coding and categorising 236 evaluations for this seven-year period helps to provide insights into the potential use of ex-post evaluations in support of future EU legislative proposals and accompanying impact assessments. Besides using our findings to enhance future policy evaluation practices, existing and future policy evaluations could be saved in a EU-wide single and central repository with open access to like-minded researchers for future similar assessments, and to interested policymakers. This could avoid duplication of efforts, reduce redundancy and repetitiveness of climate policy evaluations and allow for the sharing of lessons learnt more efficiently. Moreover, such a repository could offer interested stakeholders – including policymakers – an initial indication of the performance of climate policy in the EU and its Member States, and suggest where further capacity building for climate policy evaluation may be needed. Furthermore, the repository could offer a good basis for studying the quality of climate policy evaluations, and examining whether evaluations are in line or at odds with each other. Although the sample of evaluations covered in both meta-analyses could offer a starting point for such a repository, additional efforts and resources would be needed to collect evaluations in other Member States, and to do so on an ongoing basis. The inclusion of *ex-ante* evaluations in such a repository could be considered, so as to allow for a

comparison of whether and to what extent the expectations set out in *ex-ante* evaluations are in line with the findings of *ex-post* evaluations.

Second, addressing the evaluation criteria such as fairness and legitimacy in more Member States could enhance understanding of climate change mitigation policies across the EU. Having a good understanding on how a certain policy performed particularly according to these criteria might form the basis for more ambitious climate change mitigation policies in the future. More inclusive and participatory approaches to policy evaluations could enhance their usefulness by giving space to a variety of actors (civil society organizations, businesses, citizens, etc.) to voice their views and share their experiences when analysing policies.

Third, it is crucial and urgent to allocate sufficient resources to the coverage of relatively under-represented sectors, such as LULUCF and waste. On the sectoral coverage, the scope of evaluations is narrowed down from climate policies in general to those eligible for climate change mitigation under the UNFCCC reporting system. This study reaffirms that in the areas of mitigation policies, there is a mismatch between the sectors targeted by the majority of policy evaluations and those responsible for emissions. There is an apparent lack of *ex-post* evaluations for the agriculture, waste and forestry sectors. Thus, there is an urgent need for dedicating more resources to *ex-post* policy evaluation in the agriculture, waste, and LULUCF sectors, ideally before discussing adjusting the level of ambition in the EU's current NDC to the level required in light of the latest insight of the scientific community. When the EU explores the options to raise the level of ambition for the period beyond 2020 up to 2030 in its NDC to the Paris Agreement, the discussion should be based on sound evidence of what policies will have achieved in the EU and Member States. Commissioning studies on the performance of policies in these areas – both at the EU-wide level, and for some Member States where these sectors are responsible for a relatively large share of emissions – could help address this gap.

Despite some limitations and outstanding questions in need of further clarity this meta-analysis highlighted trends, patterns and focal areas of European evaluation practices in the area of climate change mitigation policies. Based on these outcomes, this study pointed out where the evaluation practices could be further improved and contribute to wider discussions on policy evaluations and data analysis at the European and international levels.

References

- AEA, ECOFYS, Fraunhofer and ICCS (2009). *Quantification of the Effects on Greenhouse Gas Emissions of Policies and Measures: Final Report*. ENV.C.1/SER/2007/0019. Brussels: European Commission.
- Anglmayer, I. (2016). Evaluation and ex-post impact assessment at EU level. Briefing, European Parliamentary Research Service, September.
- Ashford, N. and A. Renda (2016). *Aligning policies for low-carbon systemic innovation in Europe'*, Brussels: CEPS and i24c.
- Böbner, S., van Asselt, H., Clochard, G.-J., Alberola, E., Türk, A., Spyridaki, N.-A., Fujiwara, N., van der Gaast, W. (2017). D5.1: Report on climate change mitigation policy mapping and interaction. CARISMA

- Deliverable. <http://carisma-project.eu/Portals/0/Documents/5.1%20Report%20on%20climate%20change%20mitigation%20policy%20mapping%20and%20interaction.pdf>
- EEA (European Environment Agency) (2016). Environment and climate policy evaluation. Copenhagen: EEA.
- European Union (2004). Decision No 280/2004/EC of the European Parliament and of the Council of 11 February 2004 concerning a mechanism for monitoring Community greenhouse gas emissions and for implementing the Kyoto Protocol, [2004] OJ L49/1.
- European Union (2013). Regulation (EU) No 525/2013 of the European Parliament and of The Council of 21 May 2013 on a mechanism for monitoring and reporting greenhouse gas emissions and for reporting other information at national and Union level relevant to climate change and repealing Decision No 280/2004/EC. [2013] OJ L165/13.
- Fujiwara, N., Böbner, S., van Asselt, H., Fraser, D., Voigt, S., Spyridaki, N.-A., ten Donkelaar, M., Alberola, E., Türk, A., and K. Williges (2017). D5.2: Meta-analysis of climate change policy evaluations in the EU and Member States. CARISMA Deliverable. <http://carisma-project.eu/Portals/0/D5.2%20Meta%20analysis%20of%20climate%20change%20policy%20evaluations%20in%20the%20EU%20and%20Member%20States.pdf>
- Haug, C., Rayner, T., Jordan, A., Hildingsson, R., Stripple, J., Monni, S., Huitema, D., Massey, E., van Asselt, H. and Berkhout, F. (2010). Navigating the Dilemmas of Climate Policy in Europe: Evidence from Policy Evaluation Studies. *Climatic Change* 101(3-4), 427-445.
- Hildén, M., Jordan, A. and Rayner, T (2014). Climate policy innovation: Developing an evaluation perspective. *Environmental Politics*, 23(5). 884–905.
- Huitema, D., A. Jordan, E. Massey, T. Rayner, H. van Asselt, C. Haug, R. Hildingsson, S. Monni, and J. Stripple, (2011). The evaluation of climate policy: theory and emerging practice in Europe. *Policy Sciences*, 44. 179–198.
- Mickwitz, P. (2013). Policy evaluation. In: Jordan, A. and Adelle, C. (eds.) *Environmental Policy in the EU: Actors, Institutions and Processes*. London; New York: Routledge, 267–286.
- Mickwitz, P. and Birnbaum, M. (2009). Key insights for the design of environmental evaluations. In: M. Binbaum and P. Mickwitz (eds.) *Environmental program and policy evaluation. New Directions for Evaluation*, 105–112.
- Nilsson, M., Jordan, A., Turnpenny, J., Hertin, J., Nykvist, B., & Russel, D. (2008) 'The use and non-use of policy appraisal tools in public policy making: An analysis of three European countries and the European Union', *Policy Sciences*, 41. 335–355.
- Schoenefeld, J.J., Hildén, M. and Jordan, A.J. (2016). The challenges of monitoring national climate policy: learning lessons from the EU. *Climate Policy*. <http://dx.doi.org/10.1080/14693062.2016.1248887>
- Schoenefeld, J. and Jordan, A. (2017). Governing policy evaluation? Towards a new typology. *Evaluation*, 23(3), 274-293
- Umpfenbach, K. (2015) *Streamlining planning and reporting requirements in the EU Energy Union framework. Report for the European Climate Foundation*. Berlin: Ecologic Institute.
- Van Asselt, H., Weikmans, R. and Roberts, J. T. (2017). *Pocket Guide to Transparency under the UNFCCC*. Oxford: European Capacity-Building Initiative. http://www.eurocapacity.org/downloads/Pocket_Guide_to_Transparency_UNFCCC.pdf
- Vedung, E. (2010). Four Waves of Evaluation Diffusion. *Evaluation*, 16(3), 263-277.