The (non-)use of ex post legislative evaluations by the European Commission

Stijn van Voorst & Pieter Zwaan

To cite this article: Stijn van Voorst & Pieter Zwaan (2019) The (non-)use of ex post legislative evaluations by the European Commission, Journal of European Public Policy, 26:3, 366-385, DOI: 10.1080/13501763.2018.1449235

To link to this article: https://doi.org/10.1080/13501763.2018.1449235

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 20 Mar 2018.

Submit your article to this journal

Article views: 806

View Crossmark data
The (non-)use of ex post legislative evaluations by the European Commission

Stijn van Voorst and Pieter Zwaan

Tilburg Law School, Tilburg University, Tilburg, Netherlands; Institute for Management Research, Radboud University, Nijmegen, Netherlands

ABSTRACT

The European Commission has repeatedly emphasized that the results of ex post legislative (EPL) evaluations should be used to improve the quality of its legislative proposals. This article aims to explain the variation in such instrumental use of EPL evaluations by the Commission. Three high-quality EPL evaluations with varying levels of use were studied in-depth to assess the influence of political factors on evaluation use. The results show that, contrary to expectations, EPL evaluations may be used instrumentally even if their recommendations are opposed by important political actors in the legislative process. This article also shows that a lack of salience of the policy field to which an EPL evaluation belongs in the eyes of the Commission could, in combination with the institution’s ambition to reduce its legislative output, be a sufficient condition for the non-use of that evaluation.

KEYWORDS Better regulation; European Commission; EU legislation; evaluation use; ex post evaluation

Introduction

In its official communications, the European Commission (2015: 7, 2016: 2) has repeatedly promoted the idea of evidence-based policy: policy decisions should be based on objective information whenever possible. One important source of such information is ex post legislative (EPL) evaluations: reports that retrospectively assess the functioning of European legislation (European Commission 2015: 253). Ideally, EPL evaluations generate knowledge that allows the Commission to make informed decisions about legislative amendments (European Commission 2015: 254; Fitzpatrick 2012: 479). In both the academic literature and this article, such use of evaluations to make informed decisions about policy improvement is labelled instrumental use (Cousins and Leithwood 1986: 346).
Various academics have discussed to what extent the Commission uses evidence instrumentally in practice. Whereas some research reveals that the Commission often uses scientific evidence to improve legislative proposals (e.g., Rimkuté and Haverland 2015: 433), other studies have shown that its instrumental use of evaluations is limited due to its politicized environment and because of technical constraints (e.g., Böhling 2014: 118; Boswell 2008; Højlund 2014; Torriti 2010: 1078). Until now, such empirical research about the Commission’s use of evaluations has focused on programme evaluations (e.g., De Laat and Williams 2014; Højlund 2014) and ex-ante legislative evaluations (e.g., Radaelli 2007; Torriti 2010). Conversely, the Commission’s EPL evaluations have rarely been studied (but see Fitzpatrick 2012; Mastenbroek et al. 2016; Zwaan et al. 2016). Therefore, little is known about what factors affect the Commission’s use EPL evaluations.

This omission is unfortunate for two reasons. Firstly, due to its lack of financial and communicative tools, legislation is the Commission’s main policy instrument (Lodge 2008: 282). This makes it important to study if and how the Commission’s legislative proposals are influenced by sources of knowledge like EPL evaluations. Secondly, EPL evaluations tend to receive more attention from politicians than evaluations of other policies, since legislation affects the entire public and is usually discussed in parliament (Zwaan et al. 2016: 688). In theory, this makes it likely for the instrumental use of EPL evaluations to be affected by political conditions. For these reasons, this article answers the following research question: ‘to what extent and how do political conditions affect the European Commission’s instrumental use of EPL evaluations?’

By answering this question, this article contributes to the ongoing debate about the Commission’s nature. Originally, the Commission was perceived as a technocratic institution, aimed at impartial problem-solving (Wille 2010: 1098). Nowadays, the Commission is perceived as (also) a political actor that pursues its own preferences and acts strategically (Hartlapp et al. 2014: 1; Wille 2010: 1100). This political perspective on the Commission can be linked to a political view on evaluation use (e.g., Contandriopoulos and Brusselle 2012: 63–4; Cousins and Leithwood 1986: 347; Johnson et al. 2009: 379; Weiss 1993: 95–103). Based on these views, it can be expected that necessary conditions for the Commission’s use of EPL evaluations are that they do not contradict the preferences of the policy-makers, veto players or interest groups involved in the European legislative process.

To test these expectations, we conducted an in-depth analysis of three EPL evaluations with varying levels of instrumental use. Extensive document analysis and nineteen in-depth interviews with various actors were used to collect data about the impact of various political conditions on the Commission’s use of EPL evaluations. Technical explanations for evaluation use were
controlled for, which allowed us to better study the impact of these political conditions.

Contrary to our expectations, our results show that the absence of political opposition to specific evaluation results is not a necessary condition for use. Instead, we found a lack of salience of the evaluated policy field, combined with a commitment to limit legislative output, to be a sufficient condition for non-use. If the Commission’s political top does not prioritize a policy field, it is unlikely to follow-up on recommendations from EPL evaluations that require new legislative initiatives.

**Theoretical framework**

**Political explanations**

Since the 1970s, the evaluation literature has increasingly discussed how political conditions, next to technical ones, affect evaluation use (Johnson *et al.* 2009: 385). This literature generally argues that evaluation use is inherently political, as evaluations allocate praise or blame and may result in policy changes. Actors who feel threatened by evaluations may therefore try to prevent their use or to selectively use those results that fit their agenda (Lederman 2012: 162; Weiss 1993: 95–8).

The literature discusses several specific political conditions that affect instrumental evaluation use. A first condition is *policy-makers’ preferences* (Lederman 2012: 162; Weiss 1993: 97–8), with preferences being defined as actors’ beliefs about the feasibility and/or appropriateness of policies (Bunea 2013). Even when an evaluation recommends certain policy changes, policy-makers may oppose these changes on moral grounds (Weiss 1993: 97–8) or because they doubt their feasibility. Evaluations are often unable to change such deeply rooted policy beliefs and may therefore remain unused (Weiss 1993: 97–8).

The literature also shows that the political–institutional context of an evaluation affects its use (Cousins and Leithwood 1986: 354–5; Shulha and Cousins 1997: 196). Evaluation results are only one type of input that affects decisions about evaluation use: policy-makers are also likely to consider the position of other actors involved in the decision-making process. Since policies often result from complex negotiations between actors, policy-makers may be unwilling to reopen discussions about them when evaluations recommend to do so (Weiss 1993: 95), even when they do not object to these recommendations in principle. In particular, we expect evaluation results not to be used instrumentally if they oppose the *preferences of veto players*, as policy-makers must always reckon with the views of actors that can formally block their proposals.

Interest groups are another group of actors whose input may affect evaluation use. Such groups may have no formal veto over policy proposals, but
they can put pressure on policy-makers to ignore or implement evaluation results, either directly via lobbying or indirectly via the media. To produce a policy that satisfies a wide range of actors, policy-makers may prioritize such interest group preferences over evidence from evaluations (Shulha and Cousins 1997: 198; Weiss 1993: 95–8).

A further political condition that may affect use is the interest of politicians and civil servants to protect their financial resources (Johnson et al. 2009: 385; Weiss 1993: 95). Policy evaluations often recommend budgetary reallocations. Policy-makers may ignore such evaluation results if they view them as a threat to their own financial position.

A final relevant political condition is the media coverage of an evaluation (Weiss 1993, 95). Media coverage can influence the public opinion about the salience of issues, and issues that are high on the public agenda are likely to be acted upon. Policy-makers are therefore more likely to pay attention to and be influenced by evaluations when they have been covered by the media (Henry and Mark 2003: 303).

**Political explanations and the Commission**

Although the Commission is officially a neutral institution (Wille 2010: 1098), research increasingly shows that it (partly) functions as a political actor in reality (e.g., Hartlapp et al. 2014; Wille 2010: 1100). Concerning evaluations specifically, the Commission has been shown to ignore results from impact assessments when this was required by negotiations with the European Parliament (EP) and the Council (Torriti 2010: 1078) and to selectively use evidence from programme evaluations that legitimize its pre-existing views (Boswell 2008: 472). Based on this, we expect political considerations to also affect the Commission’s use of EPL evaluations. Below we specify how and to what extent the political conditions discussed above are relevant to explain the use of EPL evaluations by the Commission.

*Policy-makers’ preferences* are expected to matter in the context of this article. In our study, the Commission is the only decision-maker, as it has the sole right of initiative for most EU legislation and is therefore the only actor to decide about the initial follow-up of EPL evaluations (European Commission 2015: 297–8). We expect the absence of opposing preferences within the Commission to an evaluation’s recommendations to be a first necessary condition for use, since the Commission operates on the basis of a political programme (e.g., Juncker 2014) and may be unwilling to deviate from this programme when evaluation results contradict it. Since the Commission is not a unitary actor (Hartlapp et al. 2014: 2), we will consider the preferences of its two main parts involved in EPL evaluations: the directorate-general (DG) that manages the evaluation and the Commission’s political top that ultimately decides about legislative proposals.
The preferences of veto players are also expected to matter for the Commission, as there are two actors that can block its legislative proposals: the Council and the EP. The Commission may consider it useless or needlessly provocative to propose legislation that these institutions oppose, even if an evaluation recommends this (Torriti 2010: 1078). We therefore expect the absence of opposition to an evaluation’s recommendations from the EP or the Council to be a second necessary condition for use.

Interest group preferences may be especially influential in the context of this article, as the Commission actively consults such groups during most EPL evaluations (European Commission 2015: 280). Existing research shows that interest groups influence many of the Commission’s decisions, although their success depends on their resources (Bunea 2013: 567). Whereas it is common that some interest groups oppose an evaluation’s recommendations, we expect that the Commission will not implement recommendations that are opposed by all major interest groups involved in a topic. This makes the absence of such opposition a third necessary condition for evaluation use.

The interest to protect financial resources is presumably irrelevant for our study due to our focus on evaluations of legislation (i.e., non-spending activities). Media coverage is also expected to be unimportant for our study, as media coverage is generally low for EU policies – outside of some sensitive policies not discussed in this article (Princen 2011: 940). However, this expectation about media coverage will be tested in our empirical analysis.

**Technical conditions**

Besides political conditions, the literature about evaluation use also discusses several technical explanations. ‘Technical’ explanations refer to the quality of evaluation products and processes. Existing research shows that these factors influence use because policy-makers only trust evaluation results that they perceive as robust (Cousins and Leithwood 1986: 358; Johnson et al. 2009: 389; Lederman 2012: 162). Firstly, since evaluations are a form of applied research, their methodological quality matters (De Laat and William 2014: 158–60; Johnson et al. 2009: 379). Secondly, the credibility of the evaluator is important: policy-makers put more trust in evaluations published by practitioners with a sufficient reputation (Johnson et al. 2009: 379). Thirdly, an evaluation’s relevance matters: evaluations are only likely to be used if their content is required by potential users (Johnson et al. 2009: 379).

Fourthly, stakeholder involvement is important, as policy-makers can be expected to only trust evaluations that consider the views of actors directly affected by the legislation (De Laat and William 2014: 165; European Commission 2015: 280). Finally, communication quality matters: the more an evaluator stays in contact with an intended user during and after an evaluation process (preferably informally), the more likely it is that an evaluation’s findings will be
relevant for the intended user and will therefore be used (Johnson et al. 2009: 379). Thirdly, the *timeliness* of an evaluation matters, as evaluation results can only be used if they are available before important decision-making moments (De Laat and William 2014: 158). As mentioned, these technical conditions will be controlled for in this study.

**Methods**

**Case selection**

Our study is an in-depth analysis of the Commission’s use of three specific EPL evaluations. Three steps were taken to select these cases out of a dataset of 313 cases (updated version of Mastenbroek et al. 2016: 1335).

Firstly, to select evaluations for which use is likely from a technical perspective, we only considered cases that meet the criteria for a ‘good’ evaluation product and process described above. Concerning *methodological quality*, we only selected evaluation reports containing a clear operationalization and problem definition (internal validity), a representative country selection (external validity) and data triangulation (reliability). Regarding *credibility*, only evaluations by consultants who conducted at least five other EPL evaluations for the Commission were considered, as this indicates that the Commission trusts their work. Concerning *relevance*, we only selected evaluations that recommend clear legislative amendments. Regarding *stakeholder involvement*, we only selected evaluations presenting stakeholder opinions (for details about these quality criteria, see Mastenbroek et al. 2016: 1335–8).

Secondly, only evaluations published between 2008 and 2012 were considered. Evaluations from before 2008 were conducted prior to the introduction of the Commission’s evaluation procedures from 2007 (Fitzpatrick 2012: 478), meaning that any findings about such cases would be outdated. For evaluations published after 2012, it was too likely that decisions concerning their use had not been made yet.

Thirdly, after intensively scrutinizing the 12 remaining cases, three evaluations were selected. In our first case (the seed and plant propagating material (S&PM) evaluation), the Commission’s proposal was congruent with all of the evaluation’s recommendations (high level of use), in our second case (the Consumer Protection Cooperation (CPC) evaluation) the Commission’s proposal mostly followed the evaluation’s recommendations (medium level of use) and in our third selected case (the animal welfare evaluation) the Commission took no new legislative action at all, even though the evaluation recommended this (low level of use). The first two cases therefore allow us to study if the absence of opposition by influential actors is a necessary condition for use, and if so, to trace the mechanisms behind this effect. If no such causal relation is found, the comparison with the third case allows us to find other factors conditioning the Commission’s instrumental use of EPL evaluations.
The three selected evaluations offered the advantage that they were all initiated by DG Health and Food Safety (SANTE), so their organizational context was held constant. The Commission also recognized all three cases as full evaluations. Details about the three selected cases are provided in the online appendix.

**Data collection and analysis**

We collected our data via document analysis and semi-structured interviews. Commission proposals for legislative amendments (if available) were studied along with any documents leading up to them. Such documents usually included (1) an action plan based on the evaluation’s results, (2) a roadmap for legislative reform, (3) a report on stakeholder consultations, (4) an ‘incipient impact assessment’ about the expected consequences of policy options, and (5) a legislative proposal together with the final impact assessment (IA) (European Commission 2015: 297–306). The document analysis allowed us to identify which of the evaluations’ suggestions had been followed up by the Commission and which suggestions it had dropped at what moment.

Detailed explanations for these decisions were subsequently gathered via interviews, as we required open questions and follow-up questions to determine the preferences of various actors. To avoid the risk of socially desirable answers we interviewed a broad variety of respondents and guaranteed their anonymity.

In total, we conducted 19 interviews, when possible face-to-face (eight cases) and when necessary by phone (nine cases) or e-mail (two cases). For each case, we spoke to the Commission’s civil servant who had coordinated the evaluation. Regarding the animal welfare evaluation, we also spoke to the Commission’s Secretariat-General (SG), as other interviews showed it had been involved in this case. Additionally, for each case, we interviewed two respondents from different parties in the EP and two external stakeholders that had provided input for each evaluation and that represented significantly different interests (respectively small seed producers and large seed producers, consumer organizations and national consumer authorities (NCAs), and animal welfare NGOs and farmers). We did not interview the Council, as this actor did not finish discussing our second and third case at the time of writing. For each case, we also interviewed one of the consultants that conducted the evaluation.

**Operationalization**

Instrumental evaluation use, our outcome variable, refers to the consideration and implementation of an evaluation’s recommendations by its intended user to improve policies (Cousins and Leithwood 1986: 346). Therefore, we checked during the interviews if the Commission (the intended user) had considered the evaluation’s recommendations when deciding about future
policies. Furthermore, for each major legislative amendment recommended by the three evaluations, we checked via both document analysis and interviews if any subsequent legislative proposal from the Commission implemented this change. To limit our article’s scope, recommendations about legislative implementation or minor clarifications to legislation were ignored.

Concerning the political explanatory conditions, the Commission’s policy preferences were measured by asking our respondents what amendments to the evaluated legislation the Commission considered necessary before and after the evaluation was conducted. Also, for each of the major recommendations identified, we checked if it was controversial for the parts of the Commission involved in the evaluation’s follow-up (the managing DG and the Commission’s political top) and if/how this had affected the evaluation’s use.

Concerning veto player preferences and interest group preferences, respondents were asked to what extent each recommendation was in line with the views of the EP, the Council and major interest groups and how the Commission had reckoned with these views in its decisions. ‘Major interest groups’ were defined as collectives of interests (like producers and consumers) that were consulted during the evaluation. We checked the views stated by respondents with official documents when possible.

Concerning media coverage, respondents were asked if the evaluation was covered by any mainstream media up until the Commission’s decision about proposing amendments. Additionally, we analysed if the evaluations were covered by Politico/European Voice. Finally, respondents were asked if other factors had influenced the evaluation’s use.

Our assessment of the evaluations’ technical quality was checked by asking the respondents to judge the internal validity (absence of systematic errors), external validity (generalizability), reliability (absence of coincidental errors) and relevance of the final evaluation report, plus the credibility of the evaluator and the extent to which stakeholders had been involved. The timeliness of the evaluation was mapped by asking respondents if the evaluation was available to all relevant actors within the Commission when it decided about legislative amendments (Swanborn 2007: 323). Communication quality was operationalized by asking respondents how often the evaluator had in-depth contact with the Commission during the evaluation process and if informal contact was also possible (Swanborn 2007: 324).

**Results**

Below we first present the assessment of our technical conditions. We then show how each of our cases ‘scored’ on the political conditions identified above. After summarizing our results at the end of this section, we proceed with an in-depth analysis.
Technical controls

Almost all respondents who remembered the evaluations in detail confirmed that they observed high standards of validity, reliability, the credibility of the evaluator, relevance and stakeholder involvement. The sole exception was the S&PM evaluation: some respondents believed that this evaluation lacked data about small seed producers (interview 1A, 1D) and/or contained some ambiguous recommendations (interview 1A, 1B). However, these remarks only concerned some specific elements of the evaluation and other respondents did not support these criticisms (interview 1B, 1E).

Concerning timeliness, the interviews confirmed that all the evaluations were available to the Commission before it decided about legislative amendments. Regarding communication quality, in all three cases, there was frequent formal and informal contact between the Commission and the evaluator. These results confirm that the use of our three evaluations was not impeded by lacking quality.

Furthermore, none of the respondents believed that the results of the evaluations were changed significantly due to pressure from the Commission. The fact that the respondents were promised anonymity and that many of them moved to new jobs since the evaluations were completed lends some credibility to these claims, although we cannot exclude that the Commission may have subtly influenced the evaluations’ findings.

Case description 1: S&PM evaluation

The EU’s 12 directives on seed and plant propagating material (S&PM legislation) set the criteria that plant varieties must meet before they may be placed on the European market. The legislation aims to level the playing field for seed producers and to improve agricultural productivity by requiring the registration of plant varieties in national and European catalogues. This in turn requires varieties to meet standards on Distinctness, Uniformity and Stability (DUS-criteria) and, in the case of agricultural crops species, standards on Value for Cultivation and Use (VCU-criteria). Furthermore, the legislation requires national authorities to inspect the quality of individual S&PM lots (Arcadia International et al. 2008: 25–6).

As a part of the Commission’s Better Regulation Agenda, the EU’s S&PM legislation was evaluated in 2008 to suggest how its effectiveness and efficiency could be improved (Arcadia International et al. 2008: 2). Table 1 lists the evaluation’s eight recommendations for major amendments and shows which subsequent Commission documents included plans to implement them.

The S&PM evaluation represents a high level of use. Respondents from both the Commission and other organizations confirmed that the Commission took the evaluation’s findings seriously when deciding about the
Table 1. Follow-up of the recommendations of the S&PM and the CPC evaluation.

<table>
<thead>
<tr>
<th>Recommendations S&amp;PM evaluation</th>
<th>Action plan</th>
<th>Roadmap</th>
<th>Consultation report</th>
<th>Inception IA</th>
<th>Impact Assessment</th>
<th>Legislative proposal</th>
<th>Secondary legislation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Replace 12 directives with one regulation.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Make the rules for uniformity of varieties more flexible for niche markets.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Make the VCU rules evolve to adapt to types of agriculture developed for specific uses and to test varieties created by new technologies.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 4 Adapt the requirements for the marketing of seeds to defined categories. | | | | | | | N.A.
| 5 Identify links between EU seed law and food legislation. | | | | | | | |
| 6 Integrate EU plant health and seed legislation. | | | | | | | |
| 7 Grant CPVO (Commission institution) the ability to check variety denominations and the right to adopt quality requirements for DUS testing. | | | | | | | |
| 8 Reinforce provisions to inform seed users. | | | | | | | |

Recommendation CPC evaluation

| 1 Enhance the Commission’s role in the CPC network (page 17) | | | | | | | |
| 2 Expand the scope of the regulation’s annex (page 9). | | | | | | | N.A.
| 3 Give additional minimum powers of NCA’s (page 13). | | | | | | | |
| 4 Clarify NCA’s mutual obligations (page 17). | | | | | | | |
| 5 Establish procedural standards for applying NCA’s minimum powers (page 13). | | | | | | | |
| 6 Establish observatory to assist NCA’s (page 18). | | | | | | | |
| 7 Clarify the aims of the regulation (page 17). | | | | | | | |

Note: Grey cells indicate that proposals to implement the recommendations were included in the document.

The future of the S&PM legislation (interview 1A, 1B, 1E). Table 1 also shows that the Commission followed up almost all of the evaluation’s recommendations in its legislative proposal (COM(2013)262) and the preceding documents. Only recommendations 3 and 8 were ignored in some of these documents, but they were addressed through delegated acts (European Commission 2013: 5, 34).

Concerning the Commission’s policy preferences, the Commission already perceived the need to amend the S&PM legislation before the evaluation took place, as member states had notified it of various problems (like lacking harmonization) with the existing directives (interview 1A, 1B). However, according to respondents representing the Commission and the evaluator, the Commission did not have strong preferences about how the legislation should be amended (interview 1A, 1B). The Commission viewed
the legislative process as a technical matter, using the evaluation to identify potential policy improvements (interview 1A). One other respondent slightly disagreed with this and stated that the Commission had preferences in line with recommendations 1 and 7 before the evaluation was conducted, but mainly because these solutions had already been suggested by stakeholders (interview 1E).

Concerning *veto player preferences*, the EP viewed the Commission’s legislative proposal as too beneficial for large seed companies and rejected it in March 2014 (e.g., resolution A7-0112/2014). On the other hand, the Council generally supported the Commission’s views. Some countries (like France) objected to replacing twelve directives on different products with one regulation, but overall there was little controversy among the member states (interview 1A, 1B, 1E).

Concerning *interest group preferences*, large seed producers generally supported the legislative proposal (interview 1D, 1E). However, many NGOs representing small and biological seed producers criticized the proposal for how it handled recommendation 2. Most of these NGOs wanted the DUS-criteria to be abolished altogether for niche market seeds (interview 1A, 1B, 1D, 1E). Whereas the Commission’s proposal allowed such seeds to be recognized as ‘officially recognized descriptions’ to which the DUS-criteria would not apply, the NGOs viewed the criteria and procedures to apply for this exception as too demanding and opposed the existence of any compulsory registration of niche market seeds on principle (interview 1A, 1D, 1E; IFOAM EU Group 2013: 6–11).

The interviews and the media analysis showed that *media coverage* was entirely absent for this evaluation.

**Case description 2: CPC evaluation**

The EU’s CPC Regulation 2006/2004 aims to enhance the enforcement of certain European consumer protection legislation (as listed in the regulation’s annex) by increasing cooperation among NCAs. For this purpose, the regulation establishes mechanisms through which NCA’s can request each other’s assistance, including an IT platform for posting alerts. The regulation also establishes the minimum powers that national authorities must have to be able to assist each other. Furthermore, the regulation creates a European network that coordinates the activities of NCAs (the CPC network) (ICF GHK et al. 2012: 4).

Article 21a of the regulation states that it must be evaluated after five years. Accordingly, an external evaluation of the regulation was completed in 2012, which produced seven recommendations concerning major amendments (ICF GHK et al. 2012: 6–18). Table 1 lists these recommendations and shows which subsequent Commission documents included plans for their implementation.
The CPC evaluation represents a medium level of instrumental use. All respondents believe that the Commission seriously considered the evaluation’s results when deciding about possible amendments (interview 2A–2E). The legislative proposal published by the Commission in May 2016 (COM(2016)283) ignored the final two recommendations listed above, but included plans to implement the other five.

Concerning the Commission’s policy preferences, the interviews showed that recommendations 1 and 3 were longstanding priorities of the Commission, as it viewed more coordinated European action as necessary to protect consumers throughout the internal market. However, the Commission did not have strong preferences regarding the other evaluation results (interview 2C, 2E).

Concerning veto player preferences, the EP supported most of the evaluation’s recommendations as being helpful to enhance consumer protection (e.g., resolution A8-0077/2017). However, both the EP and the Council put forward amendments to remove the Commission’s right to initiate infringements (recommendation 1), as this proposal was viewed as threatening to national sovereignty. Most member states also opposed the proposed expansion of minimum powers (recommendation 3), as these powers may be difficult to handle for smaller NCAs. Furthermore, many countries opposed the content of some proposed minimum powers (like forcing infringers to compensate consumers) because legally moving these powers to their NCAs would be costly for them (interview 2B, 2C, 2E).

Concerning interest group preferences, consumer associations supported all the evaluation’s recommendations because they viewed them as beneficial for consumer protection (BEUC 2016). Business associations only objected to the proposed minimum powers to shut down websites. Our media analysis and the interviews revealed that the evaluation received almost no media coverage (interview 2B, 2C, 2D, 2E).

Case description 3: animal welfare evaluation

Legislation is one of the EU’s instruments to improve animal welfare. Various European directives protect cattle and experimental animals, for example by banning unfriendly farming methods and regulating space allowances, but most other animal types are not covered by existing EU legislation.

In 2006 the Commission published its first animal welfare strategy. In the context of this strategy, an external evaluation of the entire EU animal welfare policy was completed in 2010. Our research only concerns the part of the evaluation about legislation, which recommended to consider expanding the scope of EU animal welfare legislation to protect all animal species (GHK and ADAS UK 2010: 6).
The Commission (2012: 6) followed up on the evaluation with a second animal welfare strategy, which stated that the possibility of new animal welfare legislation should be considered in 2014. Respondents confirmed that DG SANTE took the evaluation seriously when drafting this strategy and that it would have been willing to take different decisions if the evaluation’s results had recommended this (interview 3A, 3D).

However, when DG SANTE prepared an early draft of a legislative proposal in 2014 it was informed by the SG that the proposal should wait until a new Commission would enter office in November. After this happened, the SG told DG SANTE that the existing animal welfare strategy should be fully implemented before new animal welfare legislation could be considered (even though one aspect of this strategy was considering new legislation) (interview 3A, 3D). Most respondents therefore believe the implementation of the evaluation’s recommendation to be blocked by the SG (interview 3A, 3C, 3D, 3E), although the SG states that no such decision was formally taken (interview 3F).

Concerning the Commission’s policy preferences, the animal welfare unit of DG SANTE always supported further measures to improve animal welfare, including legislation. At the top of DG SANTE and in the SG animal welfare legislation was never considered a priority, but there was little active opposition to the idea either before 2014 (interview 3A, 3D).

Concerning veto player preferences, both the interviews and various resolutions (e.g., A7-0216/2012) show that the EP strongly supports stricter animal welfare legislation (interview 3A, 3C, 3D). The Council is more divided about the topic, with countries in North(western) Europe generally supporting new legislation and some countries with much agriculture (e.g., Greece) opposing it.

Concerning interest group preferences, farmer associations opposed new animal welfare legislation because it could lead to additional costs. Animal rights groups were also sceptical about the idea of an integrated animal welfare law, as they feared it would include more self-regulation and no stricter welfare standards (interview 3D, 3E). Our interviews and media analysis revealed that there was virtually no media coverage of the evaluation.

Summary of the cross-case comparison

Table 2 summarizes the three cases and their ‘scores’ on the explanatory conditions. The CPC case has been split into two groups of recommendations that differ in their level of use; in the other cases, the level of use of the recommendations was relatively similar.

Our theoretical framework predicted that the absence of opposition to an evaluation’s recommendations from the Commission, the EP, the Council and major interest groups would be a necessary condition for use. However, as Table 2 shows this is not the case. The S&PM and CPC
evaluations are two cases where the Commission implemented respectively all and many recommendations in a legislative proposal, despite significant opposition from respectively the EP plus interest groups and the Council. The results do confirm our expectation that media coverage was absent in all cases.

To explain these findings, the next section zooms in on the steps in the follow-up process of each evaluation when specific recommendations were included in or discarded from the Commission’s plans. For the first two cases, this allows us to see why the predicted mechanisms were not triggered and if other causes can explain their outcomes instead. Our analysis of the third evaluation allows us to see if similar mechanisms can play a role in cases with low instrumental use.

Analysis

Case analysis 1: S&PM evaluation

As described above, the S&PM evaluation was entirely followed up by the Commission despite opposition from the EP and various NGOs regarding the topic of niche markets. To explain this, we must consider the Commission’s contact with these actors throughout the follow-up process. DG SANTE’s communication with the EP was mostly handled by its top-level civil servants, while the details about the evaluation’s follow-up were decided by its plant health unit. This unit received positive feedback on its plans from the member states via the comitology system, but had no contact with the EP. Accordingly, it was surprised when the proposal was rejected by the EP in 2014 (interview 1A, 1C, 1E). Therefore, the mechanism linking opposition by the EP to non-use that we predicted was not triggered.

The upcoming elections of May 2014 and a critical lobby by NGOs representing small seed producers and biological farmers both contributed to the proposal’s rejection by the EP (interview 1A, 1B, 1C, 1E). The Commission’s plant health unit had been in contact with these NGOs during the follow-up
process of the evaluation and had, as mentioned in the case description, made some concessions to their views. However, in general, the unit wished to base its proposal on the evaluation and other evidence, which it felt the NGOs did not provide. The Commission also expected that the NGOs would support the proposal in the end because it would be better for them than no change at all (interview 1A). Whereas some NGOs did indeed take this position, others opposed the amendments on principle (interview 1D). In conclusion, the Commission did not reckon with political opposition from the EP and significant interest groups because it was relatively unaware of the former and it (falsely) thought it could pacify the latter.

The Commission could have relaunched the proposal after its rejection, as some further concessions to the NGOs and the EP may have increased its chances (interview 1A, 1E). However, this option was complicated by the fact that seed legislation had no direct link to the priorities of the new Juncker Commission (the economy, human rights, migration) (Juncker 2014). Strict procedural requirements would therefore apply to any new proposal (e.g., a new IA would need to be produced), for which the plant health unit does not currently have the resources (interview 1A, 1C).

**Case analysis 2: CPC evaluation**

As described above, most of the recommendations of the CPC evaluation were implemented in a legislative proposal from the Commission despite significant opposition from especially the Council. As one respondent stated, the Commission’s proposal was ‘highly ambitious’ because it deliberately ignored objections from the member states (interview 2C). The explanation for this is that the Juncker Commission considered consumer protection a key priority to encourage the European economy (Juncker 2014: 6). When this Commission entered office it dropped many nearly completed legislative proposals to demonstrate its commitment to its Better Regulation Agenda, but the fledgling CPC proposal continued because it was considered a high priority (interview 2C, 2E).

Conversely, Table 2 also shows that the last two recommendations of the CPC evaluation were ignored by the Commission despite not going against the preferences of any influential actors. How can this be explained? Recommendation 6 (creating an observatory) was still mentioned by the Commission’s documents in mid-2014, but had been dropped by October 2015 (during which period the Juncker Commission entered office). This change was solely caused by budgetary reasons: unlike most other recommendations of EPL evaluations, establishing an observatory would cost much manpower to implement. The Juncker Commission had to reduce its civil service from the outset, and any remaining extra capacity for consumer protection was envisaged to be spent on the Commission’s increased role in the CPC
network (interview 2A, 2B, 2C). This situation appears to be a rare case where the interest to protect financial resources, which we predicted to be unimportant in our theoretical framework, does affect the use of EPL evaluations.

Recommendation 7 was not followed up because the Commission considered it to be contradictory: the evaluation first states that the objectives of the regulation must be clarified, but then states that its current objectives are ‘appropriate and relevant’ (interview 2A). Other respondents also read this recommendation in various ways, confirming its indistinctness (interview 2B, 2C, 2E).

In conclusion, the fact that some recommendations of the CPC evaluation were not followed up by the Commission is best explained by their exceptional characteristics rather than by any fundamental opposition from political actors. Conversely, the recommendations that were relatively controversial have all been followed up because the Juncker Commission viewed them as essential to its political priorities.

**Case analysis 3: animal welfare evaluation**

As was discussed above, the animal welfare evaluation’s recommendation to consider legislative changes was not followed up in the end, despite the fact that it was supported by the responsible Commission DG and the EP. Based on Table 2, an intuitive explanation for this lack of use seems to lie in the opposition of various member states and interest groups.

However, for three reasons, none of the respondents believe that this opposition was influential. Firstly, the idea of new animal welfare legislation was blocked by the SG in 2014, while member states and interest groups only seem to have lobbied about this topic at the DG-level during that time (interview 3D, 3E, 3F). Secondly, various respondents believe that an integrated animal welfare law could have been ‘sold’ to sceptical countries if it had been presented as a simplification effort, with controversial discussions about stricter welfare standards being moved to the comitology system (interview 3A, 3B, 3D). Thirdly, all interest groups state that they were much surprised when the idea of new animal welfare legislation was dropped in late 2014 (interview 3D, 3E).

So what does explain the lack of use of the animal welfare evaluation? As in the two other cases, the answer lies in the Juncker Commission’s tendency to focus on its political priorities: the economy, human rights and migration (Juncker 2014; interview 3A, 3B, 3C, 3D, 3E). To demonstrate its commitment to its Better Regulation Agenda, the Commission dropped many proposals that had no link to these topics, including the draft proposal for new animal welfare legislation (interview 3C, 3D).

In conclusion, the choice not to propose new animal welfare legislation had less to do with political preferences concerning the specific topic and more with
general shifts in the Commission’s priorities, although according to some respondents the fact that animal welfare was already considered relatively unimportant by the top of DG SANTE and the SG may also have contributed (interview 3A, 3D). The evaluation could not change this situation, as such reports are almost never read at the top of the Commission (interview 3F).

**Conclusion**

This article started with the question to what extent and how political conditions affect the European Commission’s instrumental use of EPL evaluations. Based on nineteen in-depth interviews and extensive document analysis, we traced possible reasons for variation in the levels of use of three evaluations that were all of high technical quality.

Our expectation that the absence of opposition to an evaluation’s recommendations from major political actors is a necessary condition for their use by the Commission was falsified by our findings. In our first two cases, the Commission implemented all or most of the evaluations’ recommendations, despite significant opposition from actors like the EP, interest groups and the Council. In the first case, the Commission was unaware of the EP’s opposition and falsely thought it could pacify interest groups with concessions; in the second case, the Commission considered legislative changes too important to reckon with the Council’s opposition. In our third case, opposition to the evaluation’s findings from interest groups and the Council hardly seemed to have influenced the Commission’s decision to ignore its results.

Instead, we found that a lack of salience of the policy field to which an EPL evaluation belongs in the eyes of the current (Juncker) Commission appears to be a sufficient condition for non-use. In other words, if the evaluated legislation has no direct relation to one of the Commission’s priorities, the institution is reluctant to propose amendments even when an evaluation recommends this. Our second case fits well with the Commission’s economic priorities and therefore received a legislative proposal, whereas our third case did not and therefore received no follow-up. In our first case, a legislative proposal was already dropped before Juncker entered into office, but attempts to relaunch this proposal were also hindered by the fact that seed legislation is no political priority.

What do these findings imply about the instrumental use of EPL evaluation in the Commission and in general? As our theoretical section explained, the existing literature on instrumental evaluation use (in the Commission and in general) describes various political factors which may impede such use, like the prevalence of pre-existing policy beliefs and the need to safeguard compromises. However, this existing literature pays little attention to the fact that political actors may also have a symbolic interest to reduce their policy output.
Since EPL evaluations often recommend changing legislation to improve it, they essentially request policy-makers to frequently propose legislative amendments. In the case of the Commission, such recommendations contradict its plans to propose little legislation outside of its priority fields. This contradiction leads to reduced possibilities for evaluation use.

Our findings suggest that this political interest to limit legislative proposals should be considered when studying the Commission’s instrumental use of evidence. As national executives may also commit themselves to limit their legislative output in the context of better regulation agendas, this condition may also be relevant for explaining other policy-makers’ use of EPL evaluations.

Our study has two noteworthy limitations. Firstly, due to our focus on the Commission, we did not systematically assess the wider impact of EPL evaluations on legislative outcomes. Our first case showed that even when EPL evaluations affect the Commission’s legislative proposals, they may not influence the final outcomes of EU legislative processes, as NGOs and other actors that disagree with evaluation results may still lobby against proposals based on them at the Council or EP. For future research, a more in-depth assessment of such processes would be recommended.

A second limitation lies in our case selection. Since we only studied high-quality evaluations, our conclusions may not apply to evaluations that fail to meet certain technical standards. Furthermore, because we only studied evaluations from DG SANTE the representativeness of our results could be limited, even though the selected cases covered a wide range of policies. For future research, it is therefore recommended to use a larger number of cases to study whether a lack of salience combined with a commitment to reduce legislative output is indeed a sufficient condition for non-use.

Notes

1. Other types of use often mentioned in the literature are accountability use, conceptual use and strategic use. Because these types of use may be driven by different factors than instrumental use, they are not considered in this article. For a study about the accountability use of EPL evaluations, see Zwaan et al. (2016).
2. The Commission classified three cases that met all our criteria as ‘studies’ instead of EPL evaluations. These cases were dropped to avoid unnecessary discussions about terminology.
3. The following keywords were entered in the search engine of Politico (which also shows the results for articles of European Voice) at https://www.politico.eu/: ‘seed marketing’, ‘plant seeds’ and ‘plant propagating’ (S&PM case), ‘consumer protection’ (CPC case) and ‘animal welfare’ (animal welfare case). We searched for the entire period of time between the initiation of the evaluation and either the legislative proposal or the decision not to propose legislation.
Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Stijn van Voorst is a PhD candidate at Tilburg Law School at Tilburg University.

Pieter Zwaan is an assistant professor at the Institute for Management Research at Radboud University, Nijmegen, the Netherlands.

References


