



European Union Studies Association

EU and
Public Policy

December 2015

Ex-post evaluation by the European Commission

Stijn van Voorst

Radboud University, The Netherlands





The European Commission is often said to be a frontrunner in the field of ex-post evaluation (Stern, 2009, 69). Not only has it developed its own evaluation facilities, but it has also stimulated evaluation in the member states of the EU through joint evaluations of the EU's spending programmes (Baslé, 2007, 226). While the Commission since at least the 1980s has conducted ex-post evaluations of such spending programs, evaluations of other activities are a more recent phenomenon. After being criticized for a lack of accountable government in the early 2000s, the Commission began to gradually extend its ambitions to evaluations of non-spending activities, including legislation (Fitzpatrick, 2012, 478). When the Juncker Commission entered office in 2014, its first vice president (Frans Timmermans) became responsible for the topic of ex-post evaluation. Under his guidance the Commission published new evaluation guidelines (2015), which included the ambition to conduct ex-post evaluations of entire regulatory frameworks - the so-called fitness checks. According to the new guidelines (European Commission, 2015, 259), ex-post evaluations help to improve the EU's policies, strengthen accountable government and support strategic decision-making.

These developments show that the evaluation of a broad range of policies is becoming more and more important to the Commission, at least from a rhetorical point of view. Therefore, this is an opportune moment to reflect on the topic empirically: how does the Commission's evaluation system function in reality? This article first presents some key results of academic research about the Commission's evaluation system. Based on these findings, it then offers suggestions for future research, including some possibilities to provide academic studies on ex-post evaluations with a stronger theoretical background.

To assess the performance of the Commission's system for legislative evaluation, Mastenbroek et al. (2015) conducted a meta-study of 216 ex-post evaluations of regulations and directives published by the European Commission between 2000 and 2012. They found that while the number of legislative evaluations has gradually increased over the years, the coverage of the Commission's evaluation system is far from complete. Out of all major regulations and directives from the years 2000-2012, 33% had been evaluated by the end of 2012, meaning that almost seven of out ten major legal acts had not been evaluated ex-post. The results of the meta-study also show that while more than half of the evaluations (52%) study the impact of EU legislation on society in terms of effectiveness, efficiency or side effects, the other evaluations are limited to process criteria such as transposition, implementation and enforcement. This second group of evaluations usually takes the form of a brief report to the Council and the European Parliament. The majority of evaluations concern directives (57%), with evaluations of regulations being slightly less common (43%).

Besides looking at the coverage of the Commission's ex-post legislative evaluations, Mastenbroek et al. (2015) also assessed their quality. They found that most evaluations of EU law formulate a clear problem definition (67%), contain a clear operationalization (62%) and apply a mix of different research methods (65%) - usually content analysis



combined with either interviews or surveys. About half of the evaluations clearly explain their case selection and have a response rate for their interviews or surveys of more than 50%, meaning that the external validity of the evaluations varies greatly. The reliability and methodological robustness of legislative evaluations do less well, as only 29% of the reports contain enough information to repeat the study if needed and only 15% of the evaluations offer a justification of their selected methods. The evaluations did better when it comes to usability aspects: 76% of the reports contain clear recommendations, while 64% of the reports include a short executive summary. About 60% of all legislative evaluations mention some kind of stakeholder involvement, although this is mostly limited to using stakeholders as a source of information. Only 9% of all evaluations show stakeholders being more actively involved in the evaluation process. The authors conclude that the Commission has not yet lived up to its promises when it comes to providing systematic and high-quality legislative evaluations.

Using a slightly different sample of 220 ex-post legislative evaluations, Zwaan et al. (forthcoming) addressed the question how ex-post evaluations are used after their publication. They found that the use of legislative evaluations as a tool for accountability is limited. Only 34 evaluations (16%) had been used by the European Parliament (EP) to steer the behavior of the Commission. Variance in the use of ex-post evaluations for accountability purposes turned out to be explained mainly by the political sensitivity of the legislation: evaluations of laws which were controversial in the EP turned out to be used more often to scrutinize the Commission.

Turning to the Commission's program evaluations, no single meta-study is available. This is understandable given the large number of such evaluations, which are conducted every year. However, some studies have assessed a large number of ex-post evaluations of the EU's regional policy programmes (e.g. Baslé, 2007). In general, such research shows that while programme evaluations are conducted systematically in cooperation between the Commission and the member states, they do suffer from serious methodological problems. The most prominent of these issues are a lack of transparent and reliable data sources, the problem of attributing causal effects to individual programmes and the difficulty of establishing the added value of EU subsidies compared to national policies (Baslé, 2007, 226; Bachtler & Wren, 2006, 144). Consequently, while the Commission has repeatedly used its evaluations to claim that its policies have caused increased competitiveness and growth for weak regions, such claims are disputable if the evaluations are judged by scientific standards (Bachtler & Wren, 2006, 144).

All in all, we can conclude that the coverage and quality of the Commission's ex-post evaluations varies greatly. Future research could take a more explanatory view towards these topics. For example, it is still unknown how we can explain variation in the initiation of ex-post evaluations: which policies are evaluated by the Commission and which are not? Similarly, the observed variance in methodological quality requires further analysis. From a rational perspective, we could expect the Commission to work

efficiently by focusing its evaluation activities on those policies, which are easier to evaluate. Furthermore, according to this perspective we would expect the initiation and quality of ex-post evaluations to increase as a function of evaluation capacity. However, variance in the initiation and quality of evaluations could also be explained by taking a more political perspective (Mastenbroek et al., 2015).

For this political perspective, principal-agent theory could provide a suitable theoretical framework. On the one hand, the Commission can be viewed as an agent of the Council, meaning it will want to avoid evaluations as a form of scrutiny over its activities. On the other hand, the Commission has been delegated supervisory tasks when it comes to the implementation of EU policies, so it can make use of evaluations to keep a check on the member states (Ibid.). Accordingly, we would expect the Commission to mainly invest in the evaluation of policies where the chances of non-compliance by the member states are highest, and the chances of policy reversal by Council and the EP are lowest (Mastenbroek et al., 2015).

Another interesting question is how the Commission's evaluation practice will develop in the near future. After becoming responsible for coordinating ex-post evaluation in 2009, the Commission's Secretariat General published new evaluation guidelines in May 2015. These guidelines further centralize evaluation activities, for example by standardizing the requirements for stakeholder involvement (European Commission, 2015, 264) and writing follow up reports (Ibid., 297). Until recently ex-post evaluation in the Commission was mainly a decentralized activity under the responsibility of the DGs (Stern, 2009, 69), but this might change now that the SG is becoming more actively involved. The question as to how this development will affect the functioning of the Commission's evaluation system deserves close academic scrutiny.

References

- Bachtler, J., & Wren, C. (2006). The evaluation of EU Cohesion Policy: Research questions and policy challenges. *Regional Studies*, 40(2), 143-153.
- Baslé, M. (2007). Strengths and weaknesses of European Union policy evaluation methods: Ex-post evaluation of objective 2, 1994–99. *Regional studies*, 40(2), 225–235.
- European Commission. (2015). Better Regulation Toolbox [complement to SWD(2015)111]. Brussels: European Commission.
- Fitzpatrick, T. (2012). Evaluating legislation: An alternative approach for evaluating EU internal market and services law. *Evaluation*, 18, 477-499.
- Mastenbroek, E., Voorst, S. van & Meuwese, A. (2015). Closing the regulatory cycle? A meta-evaluation of ex-post legislative evaluations by the European Commission. *Journal of European Public Policy*. Electronic publication ahead of print: <http://www.tandfonline.com/eprint/kPHKRG4wwrMCsamnGbDk/full#.Vi9L0n6rTcs> (doi:10.1080/13501763.2015.1076874).
- Stern, E. (2009). Evaluation policy in the European Union and its institutions. In: W.M.K. Trochim, M.M. Mark, & L.J. Cooksy (Eds.), *Evaluation policy and evaluation practice: New directions for evaluation* (pp. 67-85). San Fransisco, CA: Jossey-Bass.

Zwaan, P., Voorst S. van & Mastenbroek E. (forthcoming). Ex-post regulatory evaluation in the European Union Questioning the use of evaluation as instruments for accountability. *International Review of Administrative Sciences*, 82(4).