

Towards a Regulatory Cycle? The Use of Evaluative Information in Impact Assessments and Ex-post Evaluations in the European Union

Thomas van Golen and Stijn van Voorst*

As a part of its Better Regulation agenda, the European Commission increasingly stresses the link between different types of regulatory evaluations. Predictions made by Impact Assessments (IAs) could be verified during ex-post legislative evaluations, while ex-post evaluations in turn could recommend amendments to be studied in future IAs. This article combines a dataset of 309 ex-post legislative evaluations (2000-2014) and a dataset of 225 IAs of legislative updates (2003-2014) to show how many ex-post evaluations of the Commission use IAs and vice versa. This way, it explores if the Commission's rhetoric of a 'regulatory cycle' holds up in practice. Building on the literature of evaluation use, we formulate the hypotheses that the timeliness, quality and focus of the IAs and evaluations are key explanations for use. Our results show that so far only ten ex-post evaluations have used IAs of EU legislation, while thirty three IAs have used ex-post legislative evaluations. Using Fuzzy set Qualitative Comparative Analysis, we find that timeliness is a necessary condition of the use of ex-post evaluations by IAs, suggesting that for the regulatory cycle to function properly, it is crucial to complete an ex-post evaluation before an IA is launched. Future research could repeat our analysis for evaluations of non-regulatory activities or study the causal mechanisms behind our findings.

I. Introduction

During the last sixteen years, the European Commission (EC) has continuously stressed the need to improve its regulatory framework¹, an ambition which it reconfirmed most recently in its new guidelines for 'Better Regulation'.² By reducing the regulatory burden imposed on citizens and updating the legislation which remains in force, the Commission claims to promote a competitive economy and in-

crease the legitimacy of the EU.³ Impact Assessments (IAs) and ex-post legislative evaluations are two key elements of this Commission-wide agenda for Better Regulation⁴, as they are tools which can help to make legislation more 'evidence-based'.⁵ IAs are studies of the potential costs and benefits of new legislation and other major proposals⁶, while ex-post legislative evaluations study regulations and directives after they have been in effect for some time.⁷ In its communications on Better Regulation, the Commission

* Both authors are PhD-Candidates at Tilburg University Law School.

1 Communication to the Commission from Mrs. Schreyer in agreement with Mr. Kinnock and the President, "Focus on results: strengthening evaluation of Commission activities", SEC(2000)1051; Communication to the Commission from Ms Grybauskaitė in agreement with the President, "Responding to strategic needs: Reinforcing the use of evaluation", SEC(2007)213; Commission Communication, "Smart Regulation in the European Union", COM(2010)543 final; Commission Communication, "Strengthening the foundations of Smart Regulation – improving evaluation", COM(2013)686 final.

2 Commission Staff Working Document, "Better Regulation Guidelines SWD(2015)111 final.

3 Ibid., p. 4.

4 From 2010 until 2014 the Better regulation agenda was called 'Smart Regulation'. For the sake of consistency, in this article we only use the name Better Regulation, which was used in official communication before 2010 and is used again since 2015.

5 Claudio Radaelli & Anne Meuwese, "Hard questions, hard solutions: Proceduralisation through impact assessment in the EU", 33 West European Politics (2010), pp. 136 *et seq.*, at pp. 137-40; Commission SWD, "Better Regulation Guidelines", *supra* note 2, at pp. 7-9.

6 Commission Communication, "Impact Assessments", COM(2002)276 final, at p. 2.

7 Teresa Fitzpatrick, "Evaluating Legislation: An alternative approach for evaluating EU Internal Market and Services law" 18 Evaluation (2012), pp. 477 *et seq.*, at p. 478.

has increasingly stressed the need for a 'regulatory cycle' in which IAs build on the results of ex-post evaluations and vice versa to promote policy learning.⁸ For example, ex-post evaluations can test if the predictions of IAs have come true, advising the repeal of legislation which has not achieved the predicted effects. In turn, IAs can study the costs and benefits of the amendments which are meant to solve the problems exposed by ex-post evaluations.⁹

The question, however, is to what extent the rhetoric of a regulatory cycle holds up in practice. Despite the fact that the potential of linking IAs and ex-post evaluation was already recognized by the Commission more than a decade ago¹⁰, it appears that the two types of evaluation often remain unconnected in practice. In 2007, an external study of the Commission's IA system showed that only six out of twenty IAs (30%) referred to any kind of interim or ex-post evaluation, a number which includes the use of evaluations on unrelated topics.¹¹ The Impact Assessment Board (IAB) – an institution which checks the quality of IAs – stated that in 2013 only one out of six IAs (17%) made use of information from ex-post evaluations.¹²

These numbers suggest that IAs and ex-post evaluations are only loosely connected, although there is a notable lack of more systematic data on the topic from both a descriptive and an explanatory viewpoint.¹³ This article seeks to fill this gap by presenting quantitative data on the presence of a regulatory cycle in the EU. By linking a dataset of 309 ex-post

legislative evaluations with a dataset of 225 IAs of legislative updates, we are able to describe and explain how many ex-post evaluations refer to IAs and vice versa. More formally, we formulate the research question of our article as follows: how often are IAs and ex-post evaluations of EU law used in subsequent corresponding evaluative instruments and how can variance in this regard be explained? By answering this question, we also hope to provide some recommendations on how the Commission could strengthen its Better Regulation agenda.

Answering our research question does not only have a practical purpose, but also helps to improve academic knowledge. While scholars have published extensively on IAs¹⁴ and to a lesser extent on ex-post evaluations in the EU¹⁵, the connection between the two has largely remained ignored¹⁶, particularly from a quantitative viewpoint. This article helps to bridge the gap between both topics. Hopefully, the numbers we present can be a fruitful basis for future work on both IAs and ex-post evaluations in the EU.

The rest of this article is structured as follows. Section two provides background information about IAs and ex-post evaluations in the EU and how the two types of studies can be linked. In section three we present general theories on evaluation use, from which we derive a number of hypotheses about why IAs and ex-post evaluations may or may not build on each other's results. In section four our data collection, operationalization and tools for analysis are presented. Section five presents both the descriptive and

8 Commission Communication, "improving evaluation", *supra* note 1, at pp. 2-3; Commission Staff Working Document, "Better Regulation Toolbox", complementing SWD(2015)111, at pp. 288.

9 Directorate-general for Internal Market and Services, "DG MARKT Guide to Evaluating Legislation", (2008), at p. 51; Directorate-general Information Society and Media, "Evaluating Legislation and Non-Spending Interventions in the Area of Information Society and Media", (2011) at p. 17.

10 Communication for the Commission from the President and Mrs Schreyer, "Evaluation standards and good practice", COM(2002)2567 final, at p. 9.

11 The Evaluation Partnership, *Evaluation of the Commission's Impact Assessment System*, (Brussels: Secretariat-General of the European Commission 2007), at p. 86.

12 Impact Assessment Board, "Impact Assessment Board Report for 2013" available on the internet at ec.europa.eu/smart-regulation/impact/key_docs/docs/iab_report_2013_en.pdf (accessed 28-09-2015), at p. 7.

13 Stijn Smismans, "Policy Evaluation in the EU: The Challenges of Linking Ex Ante

and Ex Post Appraisal" 6 *European Journal of Risk Regulation* (2015), pp. 6 *et seq.*, at p. 19.

14 e.g. Caroline Cecot, Robert Hahn, Andrea Renda and Lorna Schrefler, "An evaluation of the quality of impact assessment in the European Union with lessons for the US and the EU" 2 *Regulation and Governance* (2008), pp. 405 *et seq.*; Fabrizio de Fransesco, Claudio M. Radaelli and Vera E. Troeger, "Implementing regulatory innovations in Europe: the case of impact assessment" 19 *Journal of European Public Policy* (2012), pp. 491 *et seq.*; Anne C.M. Meuwese, *Impact Assessment in EU Lawmaking*, (Zutphen: Wöhrmann Print Service 2008); Andrea Renda, *Impact Assessment in the EU: the state of the art and the art of the state*, (Brussels: Centre for European Policy Studies 2006); Jacopo Torriti and Ragnar Löfstedt, "The first five years of the EU Impact Assessment system: a risk economics perspective on gaps between rationale and practice" 15 *Journal of Risk Research* (2012), pp. 169 *et seq.*

15 e.g. Fitzpatrick, "Evaluating Legislation", *supra* note 7; Elliot Stern, "Evaluation policy in the European Union and its institutions" *New Directions for Evaluation* (2009), pp. 67 *et seq.*; Steven Højlund, "Evaluation use in evaluation systems – the case of the European Commission", 20 *Evaluation* (2014), pp. 428 *et seq.*

16 Smismans, "Policy Evaluation in the EU", *supra* note 13, at p. 19.

the explanatory results of our research. Section six concludes with recommendations on how the link between IA and ex-post evaluation could be improved and studied further.

II. Impact Assessment and ex-post legislative evaluation in the EU

Evaluation in the EU is primarily a decentralized activity in the European Commission. Each Directorate-General (DG) has its own evaluation-related staff and plans its own evaluation reports.¹⁷ While IAs are usually performed internally, ex-post evaluations are often outsourced to external consultants, as the Commission's staff is too small to perform these studies internally and external evaluations are believed to be more objective.¹⁸ Since 2009 a coordinating function for both IAs and ex-post evaluation lies with the Commission's Secretariat-General (SG).¹⁹

The evaluation system of the Commission has its origins in the field of programme evaluation, but from the year 2000 onwards it has increasingly been focused on other types of evaluation as well.²⁰ After receiving criticism from other EU institutions for a lack of accountability and transparency in its legislative process, the Commission launched a number of reforms in which evaluation played a key part.²¹ It became compulsory for new legislation included in the Commission's work programme and other legislation with clear social, economic or environmental impacts to have an underlying IA.²² Furthermore,

each IA was required to include a section on future monitoring and evaluation.²³ Since 2006 the quality of IAs was checked by the IAB, which was succeeded by the Regulatory Scrutiny Board in 2015.²⁴

Although systematic ex-post evaluation of EU law was also promised at the beginning of the new century, this topic only received serious attention in Commission documents from 2007 onwards.²⁵ During that year the Commission launched a communication stating that ex-post evaluation should be more integrated in the regulatory cycle to provide for Better Regulation.²⁶ In its recent policy documents on Better Regulation, the Commission increasingly stressed that closer links between IAs and ex-post evaluations are needed to increase the quality of the entire evaluation system.²⁷ The Commission's High-level group for Better Regulation has made similar remarks.²⁸ It remains unclear, however, how much of this rhetoric about a 'regulatory cycle' holds up in practice.

One particular aspect of the Commission's Better Regulation agenda is the 'evaluate first' principle, which states that an IA for a legislative amendment should always be preceded by an ex-post evaluation of the original regulation or directive.²⁹ However, in practice the decision on when to start an IA – and by extension the decision to wait for an evaluation or not – lies in the hands of the policy unit or the inter-service group responsible for the IA process.³⁰ The IAB could remark on the lack of references to ex-post evaluations when judging an IA, but this issue alone was unlikely to result in a negative opinion, which was only given in case of critical problems.³¹ It is

17 Smismans, "Policy Evaluation in the EU", *supra* note 13, at p. 19

18 *Ibid.*, at p. 22

19 *Ibid.*, at p. 7

20 Fitzpatrick, "Evaluating Legislation", *supra* note 7, at p. 478.

21 Communication to the Commission "Focus on Results", *supra* note 1, at p. 6.

22 Giacomo Luchetta, "Impact Assessment and the Policy Cycle in the EU", 4 *The European Journal of Risk Regulation* 2012, pp. 562.

23 Communication for the Commission, "Evaluation standards and good practice", *supra* note 10, at p. 7; European Commission, "Better Regulation Guidelines", *supra* note 2, at p. 49.

24 European Commission, "Decision of the President of the European Commission on the establishment of an independent Regulatory Scrutiny Board" C(2015)3263 final.

25 Fitzpatrick, "Evaluating Legislation", *supra* note 7, at p. 478.

26 Communication to the Commission, "Responding to Strategic needs", *supra* note 1, at p. 10

27 Commission Communication, "Smart Regulation", *supra* note 1, at p. 3; Commission Communication, "Improving evaluation", *supra* note 1, at p. 4; Commission SWD, "Better Regulation Toolbox", *supra* note 8, at p. 71.

28 High Level Group for Better Regulation. "Ex-post evaluation – final report", at p. 12, available on the internet at ec.europa.eu/smart-regulation/impact/docs/wg3_report_ex_post_evaluation_en.pdf (accessed 28-09-2015).

29 Commission Communication, "Smart Regulation", *supra* note 1, at p. 5.

30 Miriam Hartlapp, Julia Metz and Christian Rauh, "Linking Agenda Setting to Coordination Structures: Bureaucratic Politics inside the European Commission", 35 *Journal of European Integration* 2013, pp. 430.

31 Anne Meuwese and Suren Gomtsyan, "Regulatory scrutiny of subsidiarity and proportionality" 22 *Maastricht Journal of European and Comparative Law* (2015), pp. 483 et seq. at pp. 490-491.

possible that the link to ex-post evaluations will become more important now that the IAB has been replaced by the Regulatory Scrutiny Board, as this institution also has a formal task in judging the quality of evaluations and could thus play a role in connecting the two evaluative instruments.³² However, it remains to be seen how this will work out in practice.

Although the Commission's Better Regulation guidelines do state the need of linking IAs and ex-post evaluations, they do not go into much detail about how this can be achieved.³³ Two DGs have published guidelines for legislative evaluation which provide more information on this issue. DG MARKT states that IAs can inform evaluators which effects were expected at the time the legislation was made (the intervention logic), which could help to formulate research questions.³⁴ In turn, ex-post evaluations can suggest amendments to existing legislation which can be studied in more detail in future IAs.³⁵ DG INFSO states that IAs can also be useful to find stakeholders, as external actors consulted during an IA should ideally be consulted again during

an ex-post evaluation to see if and why their views have changed.³⁶ Furthermore, IAs can provide evaluators with background information on the topic and can notify them of potential data sources, indicators for empirical research and methodological pitfalls.³⁷

Looking beyond the Commission, the European Court of Auditors stated that ex-post evaluations often describe how policies are implemented in practice, which is useful information when drafting an IA.³⁸ In a large-scale academic study, Cecot et al. stated that the outcomes of ex-post evaluations can also be used to judge the quality of the assumptions made by IAs.³⁹ Van Gestel and Vranken⁴⁰ put this method into practice by using ex-post evaluations in the Netherlands to check the accuracy of the ex-ante assessments which the Dutch Council of State made for new legislation. The results of their research show that even though ex-ante and ex-post evaluations cannot always be compared in practice, systematically checking the outcomes of ex-ante evaluations with ex-post evaluations can be useful to strengthen ex-ante evaluations.⁴¹

Despite all these potential ways for IAs to use ex-post evaluations and vice versa, existing research on the relation between the two does not provide a very positive picture.⁴² To explain this lack of a 'regulatory cycle' despite the Commission's rhetoric on the importance of this issue, the next section presents a number of hypotheses derived from the literature on evaluation use.

III. Theories of Evaluation Use

The connection between IAs and ex-post evaluations can be conceptualised as one specific form of evaluation use.⁴³ Therefore, this chapter first presents some general explanations for evaluation use and discusses if they can also be applied to the use of IAs by ex-post evaluations and vice versa. In doing this, we focus on situations where both an IA and an ex-post evaluation about legislation actually exist - if this is not the case, this form of use is of course impossible to begin with.

Evaluation use can be defined as the way in which the results from evaluations feed back into an organisation and its policies.⁴⁴ In the literature on evaluation use there has been an extensive focus on the

32 Commission SWD, "Better Regulation Toolbox", *supra* note 8, at p. 7; European Commission, "Regulatory Scrutiny Board", *supra* note 23, at pp. 2-4.

33 Commission SWD, "Better Regulation Guidelines", *supra* note 2, at p. 30.

34 DG Markt, "Guide to Evaluating Legislation", *supra* note 9, at p. 21.

35 *Ibid.*, at p. 58.

36 DG INFSO, "Evaluating legislation", *supra* note 9, at p. 17.

37 *Ibid.*, at p. 12.

38 European Court of Auditors, *Impact assessments in the EU institutions: do they support decision-making?* (Luxembourg: European Court of Auditors 2010) Special report no. 3, at p. 40.

39 Cecot et al., "Quality of impact assessments in the European Union", *supra* note 14, at p. 409.

40 Rob van Gestel and Jan Vranken, 'Assessing the accuracy of ex ante evaluation through feedback research: A case study', in J. Verschuuren (Ed.), *The impact of legislation: A critical analysis of ex ante evaluation* (Leiden, Boston: Martinus Nijhoff 2009), pp. 199 *et seq.*, at p. 199.

41 *Ibid.*, at pp. 225-228.

42 The Evaluation Partnership, *Evaluation of the Commission's Impact Assessment System*, *supra* note 11, at p. 87; Smismans, "Policy Evaluation in the EU", *supra* note 13, at p. 13.

43 Bastiaan de Laat and Kevin Williams, 'Evaluation use within the European Commission: lessons for the Commissioner', in: Marlène Läubli Loud and John Mayne (eds.), *Enhancing Evaluation Use: Insights from Internal Evaluation Units*. (Sage: London 2014), pp. 147 *et seq.*, at p. 168.

44 *Ibid.*, at p. 150.

types of evaluation use⁴⁵ and explanations for if evaluation results are used in organisations.⁴⁶ The latter area is the focus of this article, as we are looking for the reasons why IAs may or may not use ex-post evaluations and vice versa. Three key explanations from the literature are discussed below: the timeliness of results, the quality of evaluations and the similarity of focus.⁴⁷

A first explanation is the timeliness of evaluation results: an evaluation is more likely to be used if it is published before an important decision-making moment and less likely to be used if it is presented right after.⁴⁸ Case studies about the EU's cross compliance legislation and gender research programmes have shown that timeliness is also important in the context of IAs and ex-post evaluation.⁴⁹ If an ex-post evaluation is only published while an IA of a proposed amendment is already being drafted, it is less likely the IA will take the ex-post evaluation into account.⁵⁰ In our study, we hope to find out if the conclusion of these studies holds true when analysing a larger number of cases. Therefore, we formulate the following hypotheses:

- *Hypothesis 1:* IAs conducted after an ex-post evaluation on the same legislation is published are more likely to use this ex-post evaluation than IAs conducted while an ex-post evaluation about the corresponding legislation is still being performed.

When it comes to the use of IAs by ex-post evaluations, timeliness problems are almost impossible to occur, as there is a hard requirement for an IA to be published with a legislative proposal and an ex-post evaluation only takes place once the legislation has

been in force for a couple of years. Therefore, the hypothesis formulated above only goes one way.

A second explanation for evaluation use is evaluation quality. If a report is clearly written and sound methodological choices are made, it is more likely something will be done with its results.⁵¹ Evaluations of poor quality are unlikely to be used because their results cannot be trusted and may undermine the credibility of the user.⁵² This explanation could also play a role when it comes to IAs and ex-post evaluations, as it is harder to build an IA or ex-post evaluation on earlier research in case the quality of this research is lacking. For example, even though IAs are supposed to formulate a clear intervention logic specifying causes and outcomes⁵³, this does not always happen in practice, making it harder for ex-post evaluations to refer back to them.⁵⁴ Therefore, the following hypotheses are formulated:

- *Hypothesis 2:* IAs are more likely to use ex-post evaluations which are of higher quality.

- *Hypothesis 3:* ex-post evaluations are more likely to use IAs which are of higher quality.

A third potential explanation for the (lack of) use of IAs by ex-post evaluations and vice versa lies in their difference in focus.⁵⁵ As the legal and practical requirements for IAs and ex-post evaluations differ to some extent, it can be difficult for them to build on each other's results. For example, IAs tend to be more focused on social and environmental effects and have to take future circumstances into account, while for ex-post evaluations this is not the case. In addition, IAs may be focused on one particular piece of legislation where an ex-post evaluation sometimes

45 e.g. De Laat and Williams, 'Evaluation use within the European Commission', *supra* note 40; Marlène Läubli Loud and John Mayne (eds.), *Enhancing Evaluation Use: Insights from Internal Evaluation Units*. (Sage: London 2014), at p. 3; Steven Højlund, 'Evaluation use in the organizational context – changing focus to improve theory', 20 *Evaluation* (2014), pp. 26 *et seq.*; Steven Højlund, 'Evaluation use in evaluation systems – the case of the European Commission', 20 *Evaluation* (2014), pp. 428 *et seq.*; Damien Contandriopoulos and Astrid Brousselle, 'Evaluation models and evaluation use', 18 *Evaluation* 2012, pp. 61 *et seq.*; Kim Forss, K., Claus C. Rebien, and Jerker Carlsson, 'Process use of evaluation: Types of use that Precede Lessons Learned and Feedback', 8 *Evaluation* 2002, pp. 29 *et seq.*

46 Loud and Mayne, *Enhancing Evaluation Use*, *supra* note 42, at p. 7.

47 Smismans, 'Policy Evaluation in the EU', *supra* note 13, at pp. 15-22; De Laat and Williams, 'Evaluation use within the European Commission', *supra* note 40, at pp. 158-62.

48 De Laat and Williams, 'Evaluation use within the European Commission', *supra* note 40, at pp. 158-160; Højlund 2014, 'Evaluation use in evaluation systems', *supra* note 42, at p. 429.

49 Smismans, 'Policy Evaluation in the EU', *supra* note 13, at p. 19.

50 Emanuela Bozzini and Jo Hunt, 'Bringing Evaluation into the Policy Cycle CAP Cross Compliance and the Defining and Redefining of Objectives and Indicators', 6 *European Journal of Risk Regulation* 2015, pp. 57 *et seq.*, at pp. 64-65; Lut Mergaert and Rachel Minto, 'Ex Ante and Ex Post Evaluations: Two Sides of the Same Coin? The Case of Gender Mainstreaming in EU Research Policy', 6 *European Journal of Risk Regulation* 2015, pp. 47 *et seq.*, at p. 53.

51 De Laat and Williams, 'Evaluation use within the European Commission', *supra* note 40, at p. 162.

52 *Ibid.*, at p. 162.

53 Commission SWD, 'Better Regulation Guidelines', *supra* note 2, at p. 48.

54 Smismans, 'Policy Evaluation in the EU', *supra* note 13, at p. 18; Luchetta, 'Impact Assessment and the Policy Cycle in the EU', *supra* note 22, at p. 571.

55 *Ibid.*, at pp. 17-23.

considers an entire policy field ('fitness checks') or vice versa⁵⁶ and IAs tend to be more focused on coherence rather than effectiveness.⁵⁷ As these differences in focus are expected to have a negative impacts on use⁵⁸, the following hypotheses are formulated:

- *Hypothesis 4*: the use of ex-post evaluations by IAs is affected negatively by differences in focus between the IA and the ex-post evaluation

- *Hypothesis 5*: the use of IAs by ex-post evaluations is affected negatively by differences in focus between the IA and the ex-post evaluation.

Other authors take a more institutional approach to explaining evaluation use. For example, based on a large-scale study of Swiss ex-post evaluations, Balthasar shows that internal evaluations are more likely to be used than external ones.⁵⁹ However, since the institutional setting is largely the same for all the evaluations considered in this article (for example, IAs are almost always conducted internally and ex-post evaluations are almost always conducted externally), such issues are not relevant for our purpose. The literature on evaluation use also mentions a broad dissemination of results and stakeholder involvement as key variables for explaining the use of evaluation results: the more actors know about a study, the more likely something is done with its outcomes.⁶⁰ However, since we are only talking about

the use of evaluations by other evaluations and not about use by external parties, dissemination is not relevant in the context of our study. Stakeholder involvement was included as an aspect of quality, as will be explained when our operationalization is presented in the next section.

IV. Method and Operationalization

1. Data Collection Ex-Post Evaluations

We answered our research questions with the help of two self-constructed datasets. The first dataset contains 309 ex-post evaluations of regulations and directives conducted or commissioned by the European Commission between 2000 and 2014. Since the Commission does not have one clear format for ex-post evaluations, we included reports with very different kinds of names in our dataset (the most common ones being 'evaluation', 'study', 'review', 'staff working document' and 'implementation report') as long as they have the explicit aim to study EU legislation already in force. Background studies to IAs could be included as well, as long as they fulfil this criterion. To limit the dataset to accessible legislative evaluations, we excluded evaluations focusing entirely on spending activities (even if they do have a legal basis) and five reports only available in French. Also excluded were reports that only present the data of other evaluations and studies which only concern foreign countries⁶¹ or the EU institutions⁶², as in this case there is no link to the Better Regulation agenda. In other words, the legislation needs to have an effect on citizens, organizations or member states in the EU. In case multiple evaluations by the same evaluator about the same legislation existed (e.g. annual Commission reports on a certain regulation), only the most recent one was included. Reports to the Council and the EP only presenting the results of other evaluation reports were excluded.

As the Commission's online database of evaluations is known to be incomplete⁶³, the reports were gathered from a number of sources: annual and multiannual overviews of evaluations created by the Commission (2010, 2011), the Commission's search engine for evaluations⁶⁴, Commission work programmes (2010), the EU bookshop⁶⁵, the so-called 318 reports (annexes to Commission reports on the financial regulation)⁶⁶ and lists of evaluations found on

56 *Ibid.*, at p. 18.

57 *Ibid.*, at p. 23.

58 *Supra* note 52.

59 Andreas Balthasar, "Institutional Design and Utilization of Evaluation A Contribution to a Theory of Evaluation Influence Based on Swiss Experience", 33 *Evaluation review* 2009, pp. 226 *et seq.*, at pp. 231-236.

60 De Laat and Williams, 'Evaluation use within the European Commission', *supra* note 40, at p. 167; European Policy Evaluation Consortium, Study on the use of evaluation results in the European Commission, (Brussels: EPEC 2005), at p. 61.

61 To implement this principle, we excluded all legislation initiated by the following DGs and services: DEVCO, ECHO, FPI, ENLARG.

62 With this we refer to any regulation or directive which is only binding for EU civil servants or for the legal behavior of the institutions of the EU.

63 Smismans, "Policy Evaluation in the EU", *supra* note 13, at p. 13.

64 ec.europa.eu/smart-regulation/evaluation/search/search.do, (accessed 28-09-2015).

65 [Bookshop.europa.eu](http://bookshop.europa.eu), (accessed 28-9-2015).

66 Commission Staff Working Document, "the evaluation of the Union's finances based on the results achieved", SWD(2012) 383 and Commission Staff Working Document, "the evaluation of the Union's finances based on the results achieved", COM(2013)228.

websites of the DGs. The data was checked using an existing dataset of the expertise centre Eureval, by running Google searches for evaluations of all major legislation adopted between 1996 and 2010, by searching for background documents of legislation in Eurlex and by discussing our data-gathering method with an anonymous SG employee. For a more detailed description of the data collection methods, see Mastenbroek et al.⁶⁷

2. Data Collection Impact Assessments

The second dataset used for this article contains all 225 IAs related to legislative updates which were published between the start of the IA system in 2002 and 2014. Unlike with ex-post evaluations, the Commission has a complete database of IAs available online.⁶⁸ After importing all IAs from this source (956 in total) we manually excluded the ones which are not about legislation established by the Council or the EP.⁶⁹ A first selection was made based on the titles of the IA, after which cases of doubt were checked in detail. Furthermore, IAs about legislation aimed at foreign countries or the EU's internal structure were excluded during this phase, for the same reasons we described in the previous section.⁷⁰

All 495 remaining IAs were coded for whether they relate to updates of previous legislation or to an entirely new regulation or directive. As the EU uses various words for legislative updates, we coded all cases termed as 'amending', 'recast', 'revision', 'repealing', 'simplifying' and 'supplementing' as being updates to previous legislation (225 cases in total). The two categories excluded in this way were IAs of 'new' and 'implementing' legislation (270 cases in total). The former category was excluded from our analysis because new legislation cannot be expected to build on an ex-post evaluation.⁷¹ This is not to say that IAs related to new legislation can never use data from ex-post evaluations – in fact, we encountered four cases where this was so – but in these instances the ex-post evaluation was always related to different legislation, so in this case we cannot speak of a regulatory cycle. The 'implementing' category refers to legislation which codifies an existing principle or agreement in the EU's legislative body. As these principles or agreements were not in legislation before and could not have been evaluated ex-post, we excluded them. IAs that contained both new or imple-

menting legislation as well as legislative updates were included and cases of doubt were checked manually.

As a final step, the dataset of ex-post evaluations and the dataset of IAs were cross-referenced to see for which ex-post evaluations an IA on the same legislation was available at its moment of publication and vice versa. To cross-reference the datasets we first had to link each IA to the correct regulation or directive. This was done by searching for both the number of the IA document and the number of the related Commission proposal in the European Parliament's legislative observatory.⁷²

3. Operationalization

Evaluation use, the outcome we wish to explain, was operationalized as a simple reference to an IA in the text of an ex-post evaluation and vice versa. The advantage of this method is that it allowed for a large-scale quantitative analysis, although it also means we took even very minor cases of use into account. To search for references to IAs in the ex-post evaluations, the evaluations were manually searched for the keywords 'impact ass*', 'ex ante', 'cost-benefit' and 'cost benefit', with the methodology sections of twenty report being read in detail to check if no keywords were missing. All reports where these search terms yielded results were checked manually to see if the references were indeed about IAs. To search the 225 remaining IAs for references to ex-post evaluations, we used the keywords 'evaluat*', 'ex post', 'interim', 'mid term' and 'retrospective',⁷³ with the section on procedural issues (which often states the IA's sources)

67 Ellen Mastenbroek, Stijn van Voorst and Anne Meuwese, "Naar een regelgevingcyclus? Evaluatie in de Europese Unie" 29 *Regel-Maat* 2014, pp. 212 et seq., at pp. 221-222.

68 ec.europa.eu/smart-regulation/impact/ia_carried_out/cia_2014_en.htm, (accessed 28-09-2015).

69 Such as IAs for Communications, decisions, action plans and recommendations.

70 For more details, see footnote 61 and 62 above.

71 Impact Assessment Board, "Report for 2013", *supra* note 12, at p. 7.

72 The legislative observatory can be found at www.europarl.europa.eu/oeil/search/search.do? (accessed 28-09-2015).

73 The research was done with Adobe Acrobat Reader, using an advanced search on the folders containing the IAs. Folders were divided per year.

of twenty reports being read in detail to see if no references were missing. Again, each hit was checked manually to see if there was an actual reference to an ex-post evaluation.

After finishing these initial searches, we also checked a random sample of ten IAs (out of 225) for the keywords ‘study’, ‘report’ and ‘review’, as these words are often used in the names of ex-post evaluations. These efforts yielded no additional results and since the amount of work required to search every IA for these three keywords would be disproportionate⁷⁴, we did not continue this endeavour. In case an IA did not refer to an ex-post evaluation, but we knew an ex-post evaluation was available from cross-referencing our datasets (see the previous section), we also checked any publicly available background studies to the IA for links to ex-post evaluations. This way, we found three additional references to evaluations. Some of these background studies also contained retrospective elements and were included in the sample of ex-post evaluations.

Concerning timeliness (H1), we looked at the number of months between the publication of the ex-post evaluation and the publication of the IA. According to the Commission’s official IA guidelines from 2009, conducting an IA takes about 12 months⁷⁵ so we considered the cases where the number of months was twelve or more to be ‘timely’ and the cases where the number of months was less than twelve to be ‘not timely’. Although it is not impossible for an IA to make use of an ex-post evaluation published when the IA is already under way, it is probably more difficult, as the ex-post evaluation will

not be taken into account when sources are collected at the very beginning of the IA process.⁷⁶ Therefore, we believe the threshold of twelve months is justified.

As for evaluation quality (H2 and H3), it is important to note that quality must be grounded in the subject at hand.⁷⁷ Since IAs and ex-post evaluations have different purposes to some extent – even if there are also similarities – we believe judging their quality requires different templates. For example, while ex-post evaluations should specify a clear research question, IAs always have the purpose of comparing the impacts of different policy options, meaning they have less need to explicitly state their purpose. Furthermore, while ex-post evaluations tend to contain original research, IAs often use existing data to make predictions for the future, making it hard to judge them on issues like case selection or response rate.

To judge the quality of IAs, we used an adapted version of the scorecard created by Cecot et al.⁷⁸ Since IAs are meant to compare the costs and benefits of policy options, this scorecard is focused on the quantification of alternatives. We slightly adapted the scorecard by replacing the criteria of monetized costs and benefits with two different aspects: the presence of stakeholder consultation and the presence of clear references to sources (through footnotes, a methodology section, or some other way). These changes are in line with the recent Commission standards for IAs, which emphasize stakeholder consultation and methodological soundness.⁷⁹ These adaptations to the scorecard also ensure that IAs that study subjects which can be quantified but not monetized are not put at a disadvantage. Three IAs dealing only with matters of fundamental rights and minority rights were coded as missing cases on quality, since for these issues even quantification is probably impossible. For all other types of impacts encountered, at least some quantification seemed possible.

For ex-post evaluations, we used a scorecard with eight criteria to judge their methodological quality: the presence of a clear operationalization, a clear research goal or question, an explanation of selected methods, triangulation, a replicable research design, a clear country selection, a clear case selection and a response rate of >50%. All of these criteria were weighted equally, thus creating an 8-point scale for quality (for more details, see Mastebroek et al.).⁸⁰

74 Using the three keywords ‘report’, ‘review’ and ‘study’ could generate 10.000+ hits per year. This could lead to roughly 100+ hits per IA.

75 European Commission, “Impact Assessment Guidelines”, SEC(2009)92 final, at p. 8.

76 Commission SWD, “Better Regulation Toolbox”, *supra* note 8, at p. 29.

77 T. Widmer, ‘Instruments and procedures for assessing evaluation quality: a Swiss perspective’, in: R. Schwartz & J. Mayne (Eds.), *Quality Matters: Seeking Confidence in Evaluating, Auditing and Performance Reporting* (New Brunswick: Transaction 2005), pp. 41 *et seq.*, at p. 43.

78 Cecot et al., “Quality of impact assessments in the European Union”, *supra* note 14, at p. 418

79 Commission SWD, “Better Regulation Toolbox”, *supra* note 8, at pp. 49-65.

80 Mastebroek et al., “Evaluatie in de Europese Unie”, *supra* note 63, at pp. 223-225.

Scorecard for the quality of IAs	Scorecard for the quality of ex-post evaluations
<ol style="list-style-type: none"> 1. At least some costs are stated (1p). 2. At least some costs are quantified (1p). 3. Provides point estimate or total range of costs (1p). 4. At least some benefits are stated (1p). 5. At least some benefits are quantified (1p). 6. Provides point estimate or total range of benefits (1p). 7. A measure if provided for net benefits or cost effectiveness (1p). 8. At least one alternative is considered (1p). 9. Some costs of the alternative are quantified (1p). 10. Some benefits for the alternative are quantified (1p). 11. A measure if provided for net benefits or cost effectiveness of the alternative (1p). 12. Stakeholder analysis was used (1p). 13. Information on sources is consistently provided (1p). 	<ol style="list-style-type: none"> 1. An operationalization is present. 2. A clear research aim or question is stated. 3. The methodology is explained. 4. Methodological tools are provided so that the study could be repeated if necessary. 5. Triangulation of methods is applied. 6. The selection of member states is clearly explained. 7. The selection of cases within member states is clearly explained. 8. The response rate of questionnaires and/or interviews is > 50%.
Maximum score: 13/13	Maximum score: 8/8

Table 1: Scorecard for the quality of ex-post evaluations and IAs

Table 1 summarizes the scorecards used to judge IA and evaluation quality.

Concerning the focus of the IAs and ex-post evaluations (H4 and H5), we used three indicators. First, we looked at the number of legislative acts studied by each IA and ex-post evaluation. While most reports are about a single regulation or directive, some ex-post evaluations focus on multiple pieces of legislation or even entire policy fields ('fitness checks'), which makes them potentially harder to compare with IAs. Secondly, we looked at the type of research question. Due to the nature of IAs, they are almost always focused on comparing the costs and benefits of new legislation, but for ex-post evaluations the type of research question may vary. We distinguished between evaluations which look at both costs and benefits, evaluations which only look at either costs or benefits, and ex-post evaluation which study neither costs nor benefits (e.g. pure process evaluations). Thirdly, we looked at the type of impacts on society which were studied in the IA or evaluation. Working inductively by seeing which types of impacts we found in the actual IAs and ex-post evaluations, we distinguished the following nine categories: (1) economic impacts, (2) environmental aspects, (3) employment impacts, (4) health impacts, (5) safety im-

pacts, (6) customer satisfaction impacts, (7) scientific impacts (e.g. academic output), (8) migration impacts and (9) no impacts on society. The Commission itself uses a simpler typology of three kinds of impacts (economic, environmental, and social), but we found this categorization too limited to map all the variation we observed in practice. Therefore, we distinguished between different kinds of social impacts. The final category ('no impacts on society') was used to cover evaluations only looking at transposition.

Information on each of the three indicators for focus was found by reading the introduction and methodology sections of the evaluations and the 'impact' sections of the IAs. For all three of the indicators, a comparison was then made between IAs and ex-post evaluations belonging to the same pieces of legislation to see if they were similar in focus (yes/no).

Both authors of this article coded half of the IAs for which we needed data on quality and focus. Five cases were coded together beforehand to be sure as few differences as possible would occur between the coders, and any cases of doubt were discussed immediately. The quality scores of the ex-post evaluations were taken from previous research, where intercoder reliability had already been checked and found ac-

Table 2: Operationalization

Variable	Operationalisation
IA use	0 = IA is not referred to in text of the ex-post evaluation 1 = IA is referred to in text of the ex-post evaluation
Evaluation use	0 = Ex-post evaluation is not referred to in text of the IA 1 = Ex-post evaluation is referred to in text of the IA
Timeliness	0 = Ex-Post evaluation was published 12 months or less before the IA. 1 = Ex-Post evaluation was published more than 12 months before the IA.
IA quality / evaluation quality	See table 1 above.
Object focus	0 = IA and ex-post evaluation have a different focus. 1 = IA and ex-post evaluation have a similar focus.
Problem definition focus	0 = IA and ex-post evaluation have a different focus. 1 = IA and ex-post evaluation have a similar focus.
Impact focus	0 = IA and ex-post evaluation have a different focus. 1 = IA and ex-post evaluation have a similar focus.

ceptable⁸¹, and the focus scores for the dataset of ex-post evaluations were coded by just one researcher. No cases were found where more than one ex-post evaluation was referred to in an IA or vice versa. Table 2 summarizes the operationalization described in this section.

It should be noted that a potential drawback of our study is that legislation can still be altered significantly by the Council and the EP after an IA was published by the Commission. This could make it harder for ex-post evaluations to use the results of the IA to test if all its predictions have come true. Ideally a control variable would be added to cover this condition, but unfortunately it was impossible to map which legislation was amended significantly by the EP within a reasonable timeframe. However, even in case the legislation was amended, the ex-post evaluation could still use the IA as a source for background information or background data. Therefore, this is-

sue should only affect more extensive types of use, such as testing predictions made by the IA or using it as a baseline to measure the exact effects of legislation.

4. Method of Analysis

As will appear from the results below, the number of ex-post evaluations actively using IAs and vice versa is too low to use regression analysis for our explanatory analysis.⁸² For this reason, as well as to gain a better view on which combinations of causal conditions can explain the use of IAs by ex-post evaluations and vice versa, we decided to use Fuzzy set Qualitative comparative analysis (fsQCA).⁸³ This method provides information about which (combinations of) explanatory conditions consistently generate necessary or sufficient explanations for the outcome (in the case of this article: evaluation use), based on the proportion of cases which score on both the causal conditions and the outcome.

Since most of our explanatory are coded binary, they can be used in fsQCA without problem. However, evaluation quality and IA quality are ordinal in nature and have to be transformed to vary between zero and one.⁸⁴ For such self-constructed scales

81 Mastenbroek et. al., "Evaluatie in de Europese Unie", *supra* note 63, at pp. 223-225.

82 Since we have five explanatory conditions, we would need fifty positive cases for regression analysis.

83 Charles Ragin, *Redesigning social inquiry: fuzzy sets and beyond*, (Chicago: University Press 2008), at p. 9.

84 Ragin, *Fuzzy sets and beyond*, *supra* note 77, at p. 85.

where little theoretical knowledge exists, this can simply be done by coding the lowest possible score as zero and the highest possible value as one, with the score in-between being the half-way point and the other scores being adapted proportionally.⁸⁵ As is common practice in fsQCA, we also included the negation of each condition in our analysis to check if any relations work in the opposite way of what we expected.⁸⁶

V. Results

1. Ex-post Evaluations: Descriptive Statistics

Out of the 309 ex-post evaluations in our dataset, an IA on the same legislation was available in sixty cases. Out of these sixty evaluations, only ten (17%) used the IA which was available on their topic, with a further five reports making use of an IA related to different legislation. This means fourteen cases of using an IA were found in total. Seven of these cases used the IA as a source for background information on their topic, five used data from the IA as evidence to draw conclusions from, three actively tested the predictions that the IA made concerning the costs and benefits of legislation and two used data from the IA as a baseline to measure the amount of change the legislation has caused. These categories are not mutually exclusive. One evaluation did not specify how the IA was used in any way.

Out of the sixty ex-post evaluations where an IA was available, forty cases (67%) are from 2012 or later. This makes sense given the fact that IAs were only introduced from 2003 onwards: older ex-post evaluations are often about legislation enacted before this time. All ten references which we found to IAs were also from 2012 and later, so it is too early to make any claims about developments over time.

Furthermore, eighteen ex-post evaluations referred to IAs conducted by national authorities. As EU directives must be transposed into national legislation, member states often change their laws because of EU requirements and can perform their own IAs accordingly. Fourteen of the evaluations referring to national IAs only referred to a report from the UK, which confirms this country's strong tradition in the field of IAs.⁸⁷ One report used an IA from Poland, one report used an IA from Finland, one report used

an IAs from Cyprus and Malta and one report used IAs from both the UK and Germany.

We also found ex-post evaluations which referred to IAs prospectively. Seven ex-post evaluations included an IA of a proposed amendment within their report, providing full integration of both types of evaluation. Seventeen evaluations provided suggestions for a future IA in their recommendations, proposing specific changes to the legislation of which the costs and benefits would require further study. Furthermore, four ex-post evaluations used evidence from a previous IA of a legislative amendment to support their point. In these cases, the Commission had tried to amend the legislation before, but the proposal had been rejected by the Council or the EP. However, the IA which was conducted at the time remained available to feed into future ex-post evaluations.

2. Impact Assessments: Descriptive Statistics

Out of the 225 IAs related to updating legislation, an ex-post evaluation on the legislation which was being updated was available in 51 cases. 'Available' means the ex-post evaluation was published at the time the IA was completed: ex-post evaluations published after the IA was finished were not counted here⁸⁸. For 33 out of these 51 cases (65%) we found a reference to the ex-post evaluation in the IA. 21 IAs used the ex-post evaluation as a source of background information to describe their problem, 21 IAs used information from ex-post evaluation as evidence and 20 IAs further investigated amendments suggested by an ex-post evaluation. Other forms of use were not found. Again, it should be noted that these different types of use are not mutually exclusive. Two forms of use could be identified in 17 of the 30 IAs,

85 Bruce Kogut, John Paul McDuffie, and Charles Ragin, "Prototypes and strategy: assigning causal credit using fuzzy sets". 1 *European Management Review* (2004), pp. 114 *et seq.*, at p. 123.

86 Ragin, *Fuzzy sets and beyond*, *supra* note 77, at p. 36.

87 The Evaluation Partnership, *Evaluation of the Commission's Impact Assessment System*, *supra* note 11, at p. 12.

88 Furthermore, note that some IAs refer to ex-post evaluations of individual policy programmes or action plans. Such evaluations were not counted even if the instrument they study has a legal basis.

Table 3: number of legislative IAs referring to ex-post evaluation per year

2003	0	0%
2004	0	0%
2005	0	0%
2006	0	0%
2007	0	0%
2008	7	23,3%
2009	1	3,3%
2010	0	0%
2011	9	30%
2012	2	6,7%
2013	8	26,6%
2014	6	20%
Total	33	100%

with 6 IAs containing all the three different kinds of use.

Table 3 provides an overview of the percentage of IAs referring to ex-post evaluations per year. The numbers show that there is an increase in the use of evaluations, although to this moment the 30% in 2011 is the highest number.

As mentioned above, there is a formal requirement for IAs to include a section on monitoring and evaluation.⁸⁹ We found this requirement to be applied well in practice, as only four out of the 225 IAs studied contained no information at all on a future ex-post evaluation. However, the IAs that did provide information varied greatly in their level of detail,

ranging from an extensive description of possible indicators to a simple statement about if and when an evaluation should take place.

3. Ex-post Evaluations: Explanatory Analysis

Before presenting the results of the explanatory analysis, it should be emphasized that QCA is an asymmetric method. This means that if a certain condition explains a certain outcome, this does not imply that the absence of the condition also explains the absence of the outcome.⁹⁰ The outcomes explained in this section and the next one are respectively the use of IAs by ex-post evaluation and the use of ex-post evaluations by IAs, without making any claims on how non-use of either type of evaluation can be explained.

When analysing data with fsQCA, a useful first step is to test if any of the individual causal conditions are either necessary or sufficient for the outcome.⁹¹ Table 4 below provides the explanatory analysis for the use of IAs by ex-post evaluations. As the results show, neither the quality of the IA (H3) nor the comparability of focus between the IA and the ex-post evaluation (H5) are necessary or sufficient conditions for use. In other words, the theoretical explanations which we derived from the literature do not appear to explain whether or not ex-post evaluations build on IAs of corresponding legislation.

A simple look at our data reveals similar results, as no clear patterns emerge. Out of the ten cases where we found an IA was used, six scored eight points or more on IA quality, while the other four scored five points or less. While we found that the IAs generally studied a much broader range of impacts than the ex-post evaluations (this was so in 36 out of 60 cases, or 60%), in particular when it comes to taking employment and environmental aspects into account, this issue shows no clear relationship with the fact if the IA is used. When it comes to object of study, it was the ex-post evaluations which generally had a broader scope. While 11 out of 60 cases (12%) showed an ex-post evaluation which studied multiple pieces of legislation linked to an IA which was related to just one regulation or directive, the opposite situation never occurred. However, this condition also seems unrelated to whether or not an ex-post eval-

89 Commission SWD, "Better Regulation Guidelines", *supra* note 2, at p. 49.

90 Ragin, *Fuzzy sets and beyond*, *supra* note 77, at p. 102.

91 Ragin, *Fuzzy sets and beyond*, *supra* note 77, at p. 120.

Table 4: results of QCA analysis for ex-post evaluations. Proportions ≥ 0.80 indicate a causal factor might be a necessary or sufficient condition; proportions lower than 0.80 indicate a causal factor is unlikely to be a necessary or sufficient condition. The tilde (-) represents the negation of a given variable.

Variable	Proportion cases cause > use of IA (necessary conditions)	Proportion cases use of IA > cause (sufficient conditions)
IA quality	0.60	0.21
-IA quality	0.40	0.14
Comparison of object focus	0.60	0.13
-Comparison of object focus	0.40	0.36
Comparison of problem definition	0.70	0.27
-Comparison of problem definition	0.30	0.10
Comparison of impact focus	0.50	0.23
-Comparison of impact focus	0.50	0.14

uation puts its corresponding IAs to use. The hypothesis of Smismans that the different scope of IAs and ex-post evaluations hinders the regulatory cycle is therefore not sustained by our data.⁹²

Aside from looking at the individual relations, fsQCA can also be used to analyse combinations of conditions. However, for the use of IAs by ex-post evaluations, this too yields no significant results. No single combination of conditions consistently leads to the use of IAs – in fact, consistency scores do not go above 0.28, while QCA generally requires a consistency of at least 0.80 to take a closer look at a combination of conditions.⁹³

4. Impact Assessments: Explanatory Analysis

Table 5 provides the outcome of the explanatory analysis for individual conditions which might explain the use of ex-post evaluations by IAs. As the results show, timeliness (H1) is a necessary condition

for use: if an ex-post evaluation is used, we can be almost sure that it was published at least a year before the IA. Out of the 33 IAs which used an ex-post evaluation, only four cases were untimely (12%), while for the 18 IAs which did not use the available ex-post evaluation, this was seven cases (39%). This result is in line with the finding of Bozzini and Hunt⁹⁴ and Mergaert & Minto⁹⁵ that for the regulatory cycle to function, evaluations must be available in time. However, timeliness does not seem to be a sufficient condition: even if an ex-post evaluation is published more than a year before the IA, there are still circumstances in which it is not used in the IA at all.

92 Smismans, "Policy Evaluation in the EU", *supra* note 13, at pp. 17-22.

93 Ragin, *Fuzzy sets and beyond*, *supra* note 77, at p. 125.

94 Bozzini and Hunt, "Bringing Evaluation into the Policy Cycle", *supra* note 47, at pp. 64-65.

95 Mergaert and Minto, "Ex Ante and Ex Post Evaluations" *supra* note 47, at p. 53.

Table 5: results of QCA analysis for IAs. Proportions > 0.80 indicate a causal factor might be a necessary or sufficient condition and have their level of significance provided in parenthesis. The asterisk (*) shows a result is in fact significant. Proportions lower than 0.80 indicate a causal factor is unlikely to be a necessary or sufficient condition. The tilde (-) represents the negation of a variable.

Variable	Proportion cases cause > use of ex-post evaluation (necessary conditions)	Proportion cases use of ex-post evaluation > cause (sufficient conditions)
Timeliness	0.91 (0.046)*	0.71
-Timeliness	0.09	0.27
Quality ex-post	0.57	0.67
-Quality ex-post	0.43	0.58
Comparison of object focus	0.78	0.69
-Comparison of object focus	0.22	0.47
Comparison of problem definition	0.66	0.75
-Comparison of problem definition	0.34	0.48
Comparison of impact focus	0.91 (0.046)*	0.85 (0.650)
-Comparison of impact focus	0.09	0.55

However, if timeliness occurs *in combination* with a number of other conditions, it also appears to be a sufficient condition for triggering evaluation use. The results of the analysis show one specific combination of conditions which is sufficient: the presence of a timely evaluation, the presence of a high-quality evaluation, the presence of a similar focus of IA and ex-post evaluation in terms of the number of legal acts that are studied and the similarity of the problem definition, and the absence of a similar focus in terms of the impacts which are studied. All eight IAs where this combination of conditions occurs score positively on evaluation use and the combination covers about a quarter of the total amount of cases where an IA uses an ex-post evaluation (coverage: 0.23). In other words: if an evaluation is of high quality, is produced in time and looks at the

costs and benefits of the same legislation as the IA, it is very likely it will be put to use in one way or another.

We have no logical explanation for the fact that a focus on similar impacts between IA and ex-post evaluation was absent in each of the eight cases mentioned above. After all, there is no reason why looking at *different* kinds of impacts would contribute to triggering evaluation use. But at the very least we can conclude that IAs and ex-post evaluations do not have to look at similar types of impacts for the ex-post evaluation to be used. Just like we saw in section 5.3 for the sample of ex-post evaluations, our sample of 51 IAs shows that IAs generally look at a broader range of impacts than the ex-post evaluations. This is the case for 31 out of 51 IAs (61%). In particular, IAs tend to look at environmental and employment impacts

more often than ex-post evaluations do. Nevertheless, we found eight cases in which an IA used information from a pure process evaluation, despite the fact that such reports do not factor in the costs and benefits of legislation for society. This shows it is very well possible for an IA to use an ex-post evaluation which approached its topic from a completely different angle.

6. Conclusion

This article started with the question how often IAs and ex-post evaluations of EU law are used in subsequent corresponding evaluative instruments and how variance in this regard can be explained. The Commission's increasing focus on a 'regulatory cycle' as a part of its Better Regulation agenda raises the question whether or not we can observe a link between IAs and ex-post evaluations of EU law empirically. Combining a dataset of all IAs of legislative updates with a dataset of all ex-post legislative evaluations, we have provided a first quantitative assessment of this question.

Concerning the ex-post legislative evaluations, we found that in sixty out of 309 cases an IA on the same legislation was available, but only ten evaluations actually use the IA in their report. Most of these studies used the IA as a source for background information or evidence to support their conclusions, although a small number of evaluations also tested the assumptions made during in the IA. Concerning the use of ex-post evaluations by IAs, we found that for 51 out of 225 IAs a prior ex-post evaluation on the same legislation existed, but only 33 of those IAs actually used the available ex-post evaluation in their report. This means the proportion of IAs making use of an available ex-post evaluation (65%) is much larger than the proportion of ex-post evaluations making use of an available IAs (17%). However, even for IAs there are still 35% of the cases where no use is made of an available ex-post evaluation.

One explanation for this difference could be that an IA of a legislative amendment is usually conducted right after an ex-post evaluation of the previous legislation, which means that the memory of the ex-post evaluation is still fresh. Another potential explanation is that IAs are often conducted internally, making it easier for the Commission to stimulate them to use ex-post evaluations than vice versa. A third

possible explanation is that it may be harder for ex-post evaluations to use IAs than the other way around, as IAs are often conducted before legislation is amended by the Council and the EP, meaning that their results may have little connection with the legislation which actually entered force. This problem was already recognized by the EP in their resolution on IAs back in 2011.⁹⁶ Although the Inter-institutional Agreement of 2003 stated that for substantive amendments the Council and EP should conduct their own IA⁹⁷, this principle appears not to have been applied in practice. An article in the recent proposal for a new Inter-institutional Agreement between the EU institutions shows that this problem is once again acknowledged: 'the three institutions aim to ensure that information on the impacts of the act as adopted is available, and can be used as a basis for subsequent evaluation work.'⁹⁸ To find out which of the three mechanisms stated above hinders the use of ex-post evaluations by IAs in empirical reality, more research is needed. In any case, a practical recommendation for the Commission is to require external evaluators to state whether they used the corresponding IA in their analysis and why (not).

As for our explanatory analysis, we found that timeliness is a necessary condition for the use of ex-post evaluations by IAs: an evaluation must be published at least a year before the IA, otherwise it is very unlikely to be used. The quality of the ex-post evaluation and the similarity of its focus between IA and evaluation did not turn out to be significant on their own. However, when an evaluation is timely, is of high-quality *and* looks at the costs and benefits of exactly the same legislation as the IA, we can expect it to be used. Since timeliness is so important, a practical recommendation for the Commission is to actively enforce the 'evaluate first' principle which it has emphasized in the last few years.⁹⁹ Since IAs can take

96 European Parliament resolution of 8 June 2011 on guaranteeing independent impact assessments (2010/2016(INI)), available on internet at www.europarl.europa.eu/sides/getDoc.do?pub-Ref=-//EP//NONSGML+TA+P7-TA-2011-0259+0+DOC+PDF+V0//EN, (accessed 28-09-2015).

97 Inter-institutional Agreement on better Law-making, OJ 2003 C 321/01.

98 European Commission, "Proposal for an Interinstitutional Agreement on Better Regulation" COM(2015)216 final, at p. 6.

99 *Supra* note 28; Commission SWD, "Better Regulation Toolbox", *supra* note 8, at p. 17. It should however be noted that the Commission leaves some discretionary room to ignore the 'evaluate first' principle if it is 'justified by political demands on the Commission' according to p. 256 of the Toolbox.

a year or more to conduct¹⁰⁰, it can be tempting to already launch the IA process before an ex-post evaluation is completed, but our research shows this is not a good idea if the Commission takes the idea of a 'regulatory cycle' seriously. Starting the IA process only after an ex-post evaluation is finished significantly increases the chance that the evaluation is used and the opportunity is taken to learn from how the regulation or directive has functioned in the past.

For the use of IAs by ex-post evaluations, our analysis did not reveal any (combinations of) conditions which are sufficient or necessary for evaluation use to occur. This indicates that existing explanations about evaluation use may not be adequate to explain this phenomenon. One alternative could be to look at more political models of evaluation, suggesting

that evaluation results may be used only when they are in line with preferred outcomes.¹⁰¹

Three other possibilities for future research are worth noting. In the first place, due to the Commission's rhetoric about a 'regulatory cycle', our study has been limited to legislative IAs and evaluations. Therefore, quantitative analysis of the relation between ex-ante and ex-post evaluations of spending activities still seems a fruitful field for further investigation. Secondly, our research could be repeated a few years from now, when the effects of the new Better Regulation Guidelines are in full force. Because the coupling of ex-ante and ex-post evaluative information is explicitly emphasized in the new Better Regulation Toolbox,¹⁰² it is possible that the use of evaluative instruments will increase in the near future. In the third place, while our research focused mostly on the retrospective use of IAs by ex-post evaluations and vice versa, the prospective side could also be worth studying. For example, follow-up research could see how often the plans for ex-post evaluations which are stated in IAs are executed in empirical reality. This too is an aspect of the so-called regulatory cycle.

100 *Supra* note 75.

101 Mark Bovens, Paul 't Hart and Sanneke Kuipers, 'The politics of Policy Evaluation', in: R. E. Goodin, M. Rein & M. Moran (Eds.), *The Oxford Handbook of Public Policy*, (Oxford: Oxford University Press 2008), pp. 320 *et seq.*, at p. 320.

102 *Supra* note 8, at pp. 28 and 254.