April 2025

Position Paper: The case for lean interim evaluation as a means towards adaptive EU assessment cycles

The Letta report already made the case for dynamic impact assessments that update the initial impact assessment (IA) whenever the European Parliament or the Council decide to make major changes in the draft legislation. In this position paper we extend this reasoning beyond the European decision-making process and make a case for optionally revisiting (part of the) European IAs 1 or 2 years after the legislation has entered into force. This would establish IAs as a more enduring reference point for evaluating the *actual* impact of EU legislation and could also prompt policymakers to revisit certain policy decisions if the legislation leads to outcomes that differ substantially from those projected in the initial IA. It is important to get the regulatory burdens right for all those companies and citizens that have to adhere to the legislation.

A key instrument to inform this process would be a form of lean interim evaluation. Potentially the 'reality checks' touted by the incoming Commission can fulfil this function, but as yet it is unclear what this instrument will entail. In cases where there is much uncertainty on how a piece of EU legislation will work out in practice – which we refer to as *implementation uncertainty* – such an evaluation can be planned from the onset. In practice, this consist of an ex-ante risk assessment that highlights whether draft EU legislation poses a high degree of implementation uncertainty. In such cases a so-called 'non-core review clause' could be added in the actual legislation requiring a lean interim evaluation 1 or 2 years after the legislation entered into force.

We illustrate these points with the developments surrounding the CSRD, which eventually resulted in a substantial overhaul in the first Omnibus package. From an institutional perspective, the uncertainties regarding the consequences of this act could have been dealt with in a different way from the start.

Better Regulation in conditions of uncertainty

Better Regulation is currently very much in the spotlight in the EU, as indeed it is in many Member States as well. Concerns about the competitiveness of European businesses are a strong and obvious driver for an enhanced focus on combating regulatory burden. This has not only led to a reconsideration of certain legislative measures, but also to proposals for new or repurposed instruments within the Better Regulation spectrum, such as omnibus revisions, reality checks, implementation dialogues, an enhanced competitiveness check, a stocktaking exercise of the European acquis, and more.[1] As yet, there is little information as to what most of these instruments will entail. In this paper we will argue for an element that can be a very useful addition to these practices and instruments. In a way, one could say that it is about the missing link between the *ex-ante* and the *ex-post* dimensions of scrutiny of regulatory quality. We argue

[1] See inter alia European Commission, *A simpler and faster Europe: Communication on implementation and simplification*, February 12 2025, and *A Competitiveness Compass for the EU,* COM (2025) 30.

that the continuing topicality and insightfulness of the initial European impact assessments (hereafter: IAs) can be strengthened by means of instruments of *interim evaluation*, especially – but perhaps not limited to – the cost dimension. IAs could then become a more lasting reference point, even after the European decision-making process has concluded, and in so doing more *adaptive assessment cycles* would be created, in keeping with the insights of the OECD on regulatory approaches in conditions of uncertainty.[2] We feel that the institutional mechanisms to safeguard this future-orientation of broad legislative complexes are present *in statu nascendi*, but could benefit from a more prominent profile. Companies, citizens and civil society should be able to build on the cost calculations of legislative instruments and not be faced with burdens that far exceed these estimates.

We will be illustrating this throughout with reference to the Corporate Sustainability Reporting Directive (CSRD). If there ever was a case in which doubts about regulatory burdens played a pivotal role, the CSRD appears to be it, especially in view of the complaints that were made by business and industry after its entry into force. These doubts eventually led to the first Omnibus package. It is not our intention to discuss the substantive merits of the CSRD or the need for revising it, but as a network of independent regulatory scrutiny boards we are in a good position to shed light on how the instruments of Better Regulation functioned with regard to this case.

Setting the scene: the CSRD case from an institutional perspective

The far-reaching implications of the CSRD are well known in companies and industry organisations across Europe. It has entered into force in January 2023, and has come fully into effect for a first wave of companies in the financial year 2024. The CSRD results in extensive reporting obligations for businesses. At the outset, the European IA estimated that the CSRD would impose yearly recurrent regulatory costs on businesses in a range between €1.7 and €2.3 billion, and one-off costs between €0.94 and €1.2 billion.[3] These original estimates did not account for a potential trickle-down effect, which entails that smaller companies that are 'out-of-scope' of the Directive can nevertheless be asked to report by companies they supply to. From the perspective of Better Regulation, a critical consideration is that the real costs of these reporting obligations are only becoming clear in a gradual way, as standards are drafted, delegated European legislation is enacted, national implementation choices are made, and company practices – not to mention accountancy practices – are fleshing out. Over the course of time, this limits the value of the original IA and the cost assessments made therein. In the case of the CSRD, further research at the European level has been devoted to assessing the implications of the reporting standards that have been enacted since the conclusion of the basic act and to the impact of the trickle-down effect. The most notable example is the Cost-Benefit Analysis of the first set of European Sustainability Reporting Standards (ESRS, for more information see the textbox below).[4] Similarly, a Cost-Benefit Analysis has been carried out for the more recent standards (called VSME) that specifically target SMEs.[5]

The implementation of the CSRD: ESRS, VSME and LSME
The European Commission adopted the *European Sustainability Reporting Standards* (ESRS) by means of a delegated act. These common standards are developed by the European Financial Reporting Advisory Group (EFRAG) and aim to assist companies, which are subject to the CSRD,

---

[2] See OECD, *Recommendation of the Council for Agile Regulatory Governance to Harness Innovation*, 2021; and *Practical guidance on agile regulatory governance to harness innovation*, undated.

[3] Commission Staff Working Document, Impact Assessment as regards corporate sustainability reporting, SWD(2022) 42 final, annexes, p.86. The estimated impact on businesses reflects a range because the Impact Assessment advocates a mixture of two policy options.

[4] Milieu Policy & Consulting and CEPS, 2022. *Cost-Benefit Analysis of the First Set of Draft European Sustainability Reporting Standards*.

[5] Syntesia & Prometeia, *Cost-Benefit Analysis of the VSME*, December 2024.

to comply with their sustainability reporting obligations. EFRAG was developing two new sets of sustainability reporting standards for listed SMEs and non-listed SMEs (as listed SMEs fall under the scope of the CSRD while non-listed SMEs do not): *Listed Small and Medium Sized Enterprise European Sustainability Reporting Standards (LSME ESRS)* and *Voluntary Small and Medium Enterprise European Sustainability Reporting Standard (VSME ESRS)*. The omnibus proposal, however, intends to eliminate the LSME standard.

Alongside these different cost evaluations, societal discussions on the regulatory burden of sustainability reporting – most prominently represented by the CSRD, the Corporate Sustainability Due Diligence Directive (CSDDD) and the Taxonomy Regulation – were ongoing. In the end this prompted the new European Commission to propose a radical overhaul of these Acts in its first Omnibus package. Earlier, the Commission had decided to postpone the enactment of sectoral standards to give European businesses more time to adjust. The first Omnibus package didn't come with a fully-fledged IA, but the Staff Working Document accompanying it gives a good overview of all the different consultations that were performed, and the implementation problems that European business put before the Commission.[6] Prime among these concerns was the fact that the cost of performing all these reporting duties, and in the case of the CSDDD also the duty to act on and ameliorate the reported situations, appeared to many to be much higher that the official file would have it. For instance, a Danish assessment which was also referenced in the Draghi report, concludes that the costs estimates for businesses have probably been severely underestimated in the initial IA.[7] The Danish Government, which procured this study, itself draws the conclusion that the real costs of the CSRD are five times as high as the original estimate.[8]

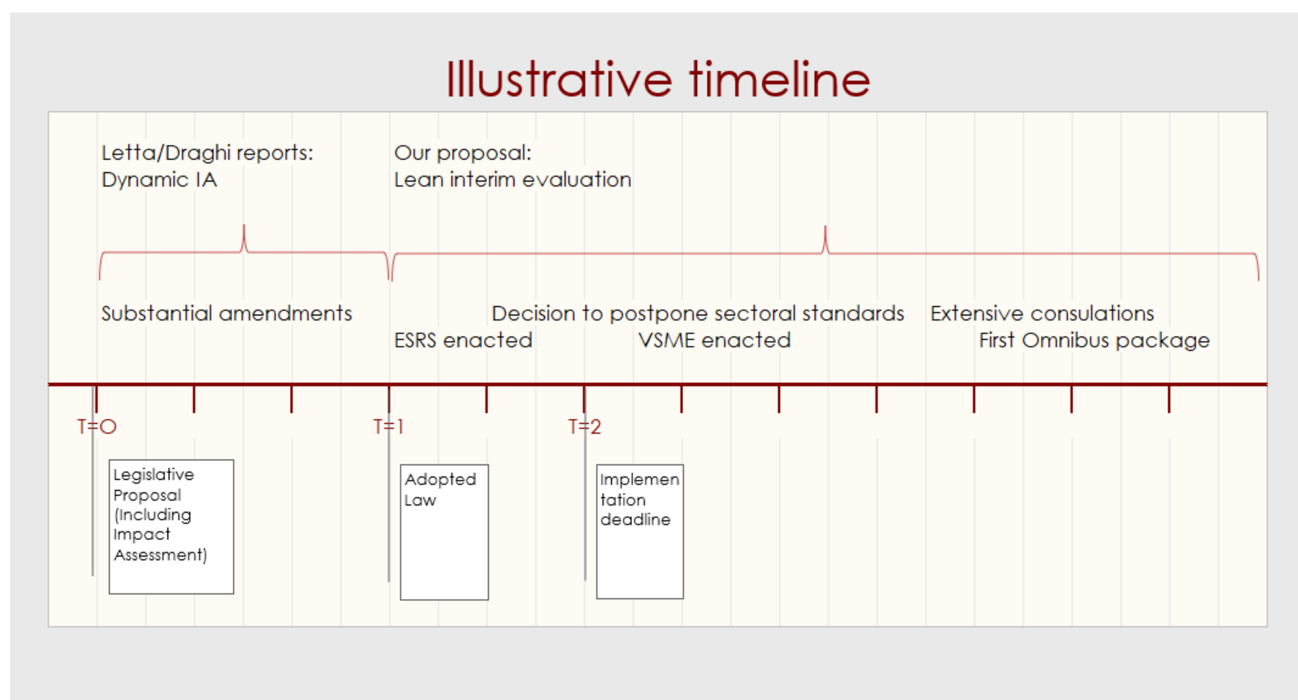
If we return our focus on the different official EU cost estimates of the CSRD, a number of things stand out that have to date not been addressed as problematic but can definitely be seen as such. Firstly, there is the fact that the different estimates are allowed to stand side by side despite their wide disparities, that is without leading to a new consolidated estimate by the Commission. In fact, the first Omnibus package doesn't draw on the original IA at all, thereby implicitly deeming it obsolete. It draws mainly on the estimates done for the ESRS, which were substantially higher than those of the IA, especially for the assurance costs. At the same time, these ESRS estimates were made by consultancy firms and weren't officially endorsed but merely "acknowledged" by EFRAG, the organisation responsible for drawing up the sustainability standards and estimating their impact. Secondly, these disparities are nowhere being addressed head-on. Notably, these also include disparities on the level of underlying assumptions. For instance, the IA had stated that the one-off costs for adjusting to the situation that limited assurance would have to be sought for the sustainability reports would be nil, while the corresponding ESRS estimates went into the billions. Thirdly, although the Standard Cost Model is universally adhered to, it does manifestly allow differences in interpretation, for instance as regards the notion of Business-as-usual (BAU) costs. This can also create large disparities between EU IAs and national follow-up exercises. Although 'BAU costs' may seem like a technocratic notion, it is actually a matter of what burdens businesses (should) consider normal – and getting it wrong can be an affront to the addressees of the norms. The extent of the BAU costs is an especially delicate question in the context of sustainability reporting, because some companies were already doing so on a voluntary basis – guided by inter alia OECD principles – and many more would probably have followed on the basis of the 'reputation mechanism', i.e. due to societal pressures. Fourthly, there is the very important issue of the indirect effects of the CSRD, notably for SME's that are in the value chains of larger

[6] SWD (2025) 80 final.

[7] COWI, AMVAB – DIREKTIV OM VIRKSOMHEDERS BÆREDYGTIGHEDSRAPPORTERING (CSRD), March 2024. Also referenced (without source) in part B of the Draghi report, p. 318.

[8] Ministry of Industry, Business and Financial Affairs, *Danish input to how to reduce administrative burdens*, February 2024.

companies and would be required to report to them. This 'trickle-down effect' was only fully quantified by the cost benefit analysis of the VSME in December 2024, and turned out to be especially costly. The IA had refrained from even making an approximation, stating that this could not be done on the basis of the information that was available at that time. At the very least one could state that the EU lawmaking Institutions could not draw on complete information when they enacted the CSRD.



Ensuring the continued topicality of IAs

An IA's primary purpose of infusing evidence into the original European decision-making process is one thing, but its continued value as a reference point for ascertaining the impact of the measures adopted quite another. Our discussion of the different approximations of the costs of the CSRD should have made clear that the value of the original IA has become very tenuous in the course of developments *during* and *after* the initial European decision-making process. In general, a lot of attention is being paid to the problem that IAs are often not updated during the European even though substantial amendments are being made, notwithstanding that the European Parliament and the Council agreed to do so in the Interinstitutional Agreement on Better lawmaking. In this context, the Letta-report made the call for *dynamic impact assessments* to update the initial IA after amendments have been made by the Council and the European Parliament.[9] This topic has been at the forefront of attention for many years, and perhaps the upcoming revision of the Interinstitutional Agreement will create conditions for real improvement in this respect. An observation can be added to the existing discussion on these matters, and that is that even if the Council and European Parliament continue to experience problems in making good on their promises, the Commission could decide that a revision of the IA is called for at the moment of enactment of the instrument. This would be especially important in cases like the CSRD file, where a large portion of the regulatory burdens emanate from standards and delegated acts that were unknown at the time of the drafting of the basic instrument. Be that as it may, the purpose of this paper is to draw attention to what happens *after* the European decision-making

[9] Letta Report, *Much more than a market*, 2024, p.124.

on the basic act, and in particular to the question whether the IA could continue to serve as a reference point. As the case of the CSRD shows, this is not a given.

To be sure, drafting an IA in conditions of uncertainty such as is the case with the CSRD is no mean feat. Many extensive and complex EU laws pose serious problems in determining the amount of costs and benefits with some level of certainty, which can be characterised as *implementation uncertainty*. As said, in situations of uncertainty the emphasis should be on the adaptiveness of the assessment and feedback cycle. In connection with IAs, this is important for several reasons. Firstly, an IA is a generally thorough piece of *policy analysis*, by means of which one can later elucidate what unforeseen effects have come to the fore after the rules have entered into force. This in theory should be an important *qualitative* policy feedback loop, that we do not always see functioning in practice. Secondly, an IA should serve as the starting point of a *quantitative* feedback mechanism as well. When the calculations concerning the costs and benefits turn out to have been unrealistic in hindsight, this should be an important impetus for political reconsideration. Thirdly, the calculations in European IAs factually already have an 'afterlife' in the national domain, even if the calculations are somewhat obsolete because of the amendments made in the European decision-making process. By this we mean that European IAs are often unquestioningly repurposed for national policy, without taking into account the 'trilogues and tribulations' of European decision-making that can in theory produce a policy option that the original IA expressly rejected and therefore did not calculate the impacts for.

Interim evaluations in the present policy cycle of the EU

In our view, the missing link between *ex-ante* and *ex-post,* or in concrete terms in keeping IAs topical, could be formed by the instrument of interim evaluation (sometimes called *ex durante* evaluation). Interim evaluations can be used specifically to revisit cost calculations of legal acts when there are signs that the ensuing implementation issues have rendered the original calculations obsolete. In fact, uncertainties accompanying the basic act can be addressed from the very start on the basis of some kind of risk assessment, and interim evaluations put in place in a forward-looking fashion. To quote the OECD, "The type of ex-post review, and its timing or "triggers", are generally best determined at the time regulations are being made."[10] In turn, these evaluations should lead to official consolidated cost calculations that update the original file in an authoritative way.

Such interim evaluations need not be fully-fledged evaluations with all their methodological rigour and breadth of scope. There is no need for a heavy instrument that revisits the entirety of the IA. Rather, it should be a lean instrument that focuses on issues of practicability, with optional further steps with regard to a revised assessment of real costs. Such 'lighter' evaluation can supply the European Commission with the necessary information to initiate legislative revision in cases where there are concerns for the workability of the legislation for businesses or the costs show to be much higher than expected in the initial impact assessment.

There are beginnings of such an instrument in the European policy cycle, but they are as yet not structurally connected with revisiting the impacts of a given measure in conditions of initial uncertainty. If these existing instruments would be repurposed to serve the goals we are propagating here, particular attention should be given to the *way they are triggered* and what the *nature of the analysis* is. As for the trigger mechanism, it makes sense to primarily think of the inclusion of special review clauses in proposed acts that entail a lot of uncertainty. It would fall on the Commission services to give the first push for this, but one can also envisage a role for the Regulatory Scrutiny Board (RSB). If the draft IA has not already done so, it would make sense that the RSB would point out the high level of uncertainty in their scrutiny of the calculations of

[10] OECD, *Reviewing the Stock of Regulation*, 2020, pp.23-24.

costs (and benefits) in the IA Reports. That would de facto mean a 'comply or explain' obligation for the subsequent law drafting within the Commission. At the end of the line, a conclusion on the need for interim evaluation needs to be embodied in the piece of legislation itself, so the European Parliament and Council have to sign off on it as co-legislators.

As for the *nature of the analysis* in such interim evaluations, we first have to delve deeper into the existing structures of evaluation in the European landscape. A fully fledged evaluation covers more than just a factual assessment of the effects; it ascertains why the effects have taken place, what part can be attributed to the EU intervention and to which extent the change matches the original expectations/predictions.[11] The evaluation clauses specify how, when and which policies will be evaluated.[12] In this context the distinction between core and non-core review clauses is useful.[13] The clauses can instruct that a full evaluation has to take place years after the moment the Directive or Regulation enters into force, in which case we talk about a core review clause. In contrast, non-core review clauses do not call for a fully-fledged evaluation but instead for a lighter kind of evaluation 1 to 3 years after the Regulation or Directive entered into force. The call for 'reality checks', as mentioned in the Mission Letter from the President of the European Commission to the Commissioner-Designate for Economy and Productivity, can in this context be seen as new type of lighter evaluation that aims to identify hurdles for stakeholders when implementing the new rules of a recently enacted legislative act.[14] A first search for evaluations in the EU policy cycle shows that approximately 60% of the adopted EU legislative acts in the Ordinary Legislative Procedure include a review clause.[15] The data shows that 93% of the review clauses in adopted Directives involve a core review clause while for Regulations this is 85%, highlighting the importance in the EU allocated to fully fledged evaluations a number of years after the legislative act entered into force.[16]

Although fully fledged evaluations seem to be well-established, it is less clear whether non-core review clauses and the use of lighter interim evaluations have reached their full potential.[17] In fact, the inclusion of a non-core review clause in the CSRD to request the European Commission to assess after 1 to 3 years the impact on businesses of the obligation to report on sustainability issues would have provided the opportunity to update the estimates in the initial impact assessment and, if needed, refine the CSRD. What ensued instead was a slightly chaotic societal debate that eventually led the Commission to propose a complete overhaul – which poses major issues in terms of legal certainty – without the road to that point being especially transparent.

Experiences from the Netherlands and Germany

As the European landscape is currently dominated by fully fledged evaluations, we would like to turn to some national experiences for inspiration on how to perform more restricted evaluations in a meaningful way. We have examined the experiences with the Dutch Implementation test and the German Reality Check as Better Regulation tools that can be used when there is a high degree of uncertainty with respect to the consequences of a proposal. The Dutch instrument can certainly be viewed as an interim evaluation tool which is carried out relatively soon after the moment the new legislation has come into effect. The rationale for this type of evaluation is to get an early

---

[11] Better Regulation Toolbox, tool 45.

[12] Better Regulation Toolbox, tool 44, gives more details about the way the evaluation clauses might be carried out.

[13] European Parliamentary Research Service, *Review clauses in EU legislation adopted during the first half of the ninth parliamentary term (2019-2024)*, 2022, p.9.

[14] European Commission, Mission Letter to Valdis Dombrovskis - Commissioner-designate for Economy and Productivity, September 2024.

[15] It concerns adopted legislative acts between 1 October 2019 and 31 December 2021.

[16] European Parliamentary Research Service, op. cit., pp. 15-17.

[17] We are not concerned with interim evaluations of European funds here. These so called midterm reviews are very common, but are of a different nature than non-core reviews of substantive legislation.

warning about unexpected effects and problems after the implementation of the new law. This will allow policy officials to address the need for amendments as they arise. As for the German instrument, its precise profile and use case is yet to be developed. But it deserves mention because the notion of reality checks has been taken up in the EU, and it could move in the direction of a tool for interim evaluation.

The Dutch Implementation Test ('Invoeringstoets') is described as a light study of the impact of new legislation in practice, with a special focus on target groups and executive agencies. This test is carried out at the first possible moment that something useful about the impact in practice can be determined. That is usually between 1 and 2 years after enactment of the rules. The test explicitly does not focus on effectiveness but on bottlenecks and problems in practice, and those can also be side effects of the legislation. Such problems can influence the effectiveness of the law, but uncovering that is not the direct aim of the Implementation test. However, whenever the test shows that firms complain about having to invest more effort to comply with the obligations than predicted ex-ante, a study can be carried out to revisit the ex-ante assessment. There is about two years of experience with Implementation Tests in the Netherlands. In this period about 45 tests were carried out. The first results are positive in the sense that it served as an effective early warning system. In most tests one or more bottlenecks were found. Criticisms of the tests in the Netherlands state that it is not self-evident when such Tests are being held. For example, politically sensitive legislation on energy transition and environmental protection has not been subjected to an Implementation Test. [18]

The German practice of Reality Checks is another example of an instrument that seeks to identify implementation problems 'on the ground' at a relatively early stage. This practice has probably already served as an inspiration for the European domain, seen that Reality Checks are prominent among the new instruments of Better Regulation that the Commission has launched. The emphasis – both in the German and the EU context – appears to be on conversations with companies to get at the practical problems they are experiencing. It is in that sense definitely meant as an early warning system that can feed back into a possible policy revision at a much earlier time than a normal evaluation would. However, the precise characteristics of the Reality Checks have not yet been well established, in Germany[19] let alone at the EU level. To serve the purposes propagated here, a European Reality Check would need to go beyond gathering experiences on the ground, but also seek to quantify them in a systematic and overarching way by means of extrapolations etc.

The case for using lean interim evaluations to deal with implementation uncertainty

Given the prominence of burden reduction in the current European landscape, it would appear fitting that more effort is made on keeping the evidence base of legislation up to date and not abandoning that effort once the measure has come into effect. Of course there is a lot of monitoring going on in Brussels, but if the data that this yields only comes to the surface when a full evaluation is performed (say) seven years later, this is much too late in many respects. Lean instruments of interim evaluation are well suited for the task of providing early warnings that the underpinnings of a legal act – be it assumptions or cost calculations – are being overwritten by the facts on the ground. Unsound cost calculations or assumptions can have a damaging effect on companies, citizens and civil society organisations who stake out ways to deal with regulation on the basis of the official information provided.

---

[18] Arno F.A. Korsten, *Early warning evaluation en de invoeringstoets* (in Dutch), Beleidsonderzoek Online 2023.

[19] Cf. National Regulatory Control Council (Germany), Annual Report 2024, *Good legislation. Digital administration. Less bureaucracy*.

In sum, our argument towards more adaptive assessment cycles is threefold. First, the function of an IA as a lasting reference point should be strengthened in order for it to remain valuable in the monitoring of real effects of a European measure. Second, the way to do so could be to develop lean interim evaluation instruments by which an early warning regarding actual practical impacts and costs can be generated. Such 'lighter' evaluation would in first instance involve an analysis of problems related to practicability of the legislative act for businesses and/or citizens, and be followed up upon by an update of the cost assessment. Third, these instruments need to feed back into and contribute to closing the European policy cycle. Thus, a structure for more adaptive assessment cycles can be created that goes beyond the European decision-making process both in scope and in time, in that it can prove valuable for national follow up exercises as well. The whole evidence base of a legal instrument is kept dynamic in this way, which should also help towards avoiding the need for complete overhaul such as the first Omnibus package. However, for later evidential updates to be able to feed into the policy process, there needs to be a degree of coherence and consolidation of the methods. For instance, the lack of *any* approximation of the trickle down effects of the CSRD in the IA appears in hindsight to have been a large omission. The same can be said of the divergent ways to interpret what BAU costs are.

We want to stress that there is also a great responsibility for Member States to contribute to such follow up assessments. The incentive should however be clear, as this provides a structured way to bring national experiences with implementation back to the European table, and the exercise is notably not performed in the register of compliance but of feasibility and learning.