Disclaimer:

The following text corresponds to the article as written in the original language, before being translated into French at the request of the RFAP. The present version, therefore, is the sole responsibility of the author, the editorial validation concerns only the version published in French. Any quotation or bibliographic reference to this article must refer to the French version, as published in the RFAP, rather than to the version in the original language.

Avertissement:

Le texte reproduit ci-après correspond à l’article tel qu’il a d’abord été rédigé dans la langue originale, avant sa traduction en français à l’initiative de la Revue. La présente version relève donc de la seule responsabilité de l’auteur, la validation éditoriale de l’article par la Revue portant uniquement sur la version publiée en langue française. Toute citation ou référence bibliographique devra impérativement viser le texte en langue française tel que publié dans la Revue, et non le texte dans la langue originale.
THE APPRAISAL OF POLICY APPRAISAL: 
LEARNING ABOUT THE QUALITY OF IMPACT ASSESSMENT

Claire A. Dunlop
Oliver Fritsch
Claudio M. Radaelli

Abstract
What do governments, international organizations and stakeholders mean when they say that proposals for new regulation should be systematically appraised? And do regulators really use the results of appraisal? In this article, we consider two dimensions of policy appraisal: the breadth and scope of the empirical analysis, and the utilization of impact assessment. We use these two dimensions to produce an explanatory typology with four types. The types enable us to review the literature systematically, exposing gaps as well as documenting the results. In the final part of the article we build hypotheses that link quality of analysis and utilization, thus showing how future research may become less descriptive and more inclined to test explicit hypotheses.
1. Introduction

Robust, evidence-based appraisal of proposed regulation is a fundamental aim of the so-called ‘smart regulation’ agenda pursued by governments, the European Union (EU) and international bodies such as the Organization for Economic Co-operation and Development (OECD). In the past decade, as impact assessment (IA) has spread across most OECD countries, the debate on IA has moved from issues of adoption (Radaelli 2005; Staronová 2010; De Francesco, Radaelli and Troeger 2012) to those of usage and utilization. For policy makers, audit bodies and social scientists alike the key preoccupation concerns how to get results from this policy instrument. The EU in particular has added a multi-level dimension to these challenges, because the domestic and EU-level appraisal systems have several points of contact. For example, EU legislation is first appraised by the European Commission, and then it is examined with IA tools at the stage of implementation into national legal systems. This explains why the topic of ‘quality of IA’ – or the appraisal of policy appraisal – is so popular today across Europe and the OECD countries.

Policy makers have expressed interest (in various fora; such as the EU and the OECD) on quality benchmarks in the adoption and implementation of IA. There are various motivations for this. The first ambition is to manage regulatory systems. To illustrate, indicators on the quality of IA help government-wide regulatory oversight bodies map progress on their regulatory reform agenda and stimulate discussion with departments and regulatory agencies. Linked to this, information on IA alerts the officers responsible for the overall regulatory reform agenda on where exactly progress is slower. Indeed, indicators and measures should point to the remedies or program changes that would improve performance. The second push is communication. Information on the performance of IA and regulation is a key element of a communication strategy at different stages of regulatory reform. This is particularly important in restoring trust in the business environment and in the regulatory systems, among firms as well as citizens. Finally, IA appraisals are carried out to promote accountability. Departments and agencies are accountable to a variety of political institutions (in primis, the Parliament) and ‘social constituencies’. Clearly, the most important end-users of regulatory reform are ministers, parliamentarians, and citizens. Gathering information on the quality of IAs is a way for these actors to hold the bureaucracy to account.

However, what does quality actually mean in this context? How can it be measured? And, what are the best strategies to explain it, and make use of this information in managing regulatory reform? We address these questions by reviewing key studies in the literature appraising IA. Section 2 explores different ways in which quality can be conceptualised and introduces a four-fold typology. This typology is then used in Section 3 to structure our selective review of the literature – with examples drawn from the United Kingdom (UK), United States (US) and the EU. Section 4 draws
the article to a conclusion, suggesting three main findings and building hypotheses for further research.

2. The concept of quality

Notions of IA quality, and their measurement, are intrinsically related to the standard benchmarks of regulatory quality. Radaelli and De Francesco (2007) revise the standard benchmark of regulatory quality, such as efficiency, access to the regulatory system, transparency, use of evidence, and involvement of stakeholders via participation. But when we talk about IA, we are not immediately interested in the quality of rules, but in the quality of the process that has produced a given regulation. Radaelli and Fritsch (2012) suggest that regulatory quality can be examined along the whole life cycle of regulation:

- **Input** Regulatory quality in input addresses what happens at the outset of the regulatory life cycle with design activities. Specifically, the input dimension includes the formal adoption of policy appraisal, the provision of training events on the economic analysis of regulation and consultation, as well as the establishment of a regulatory oversight body that scrutinises the appraisals produced.

- Moving forward in the regulatory life cycle, the next dimension of quality concerns **process**; connecting the inputs and producing results. Typically, processes define the scope, and extent, of tools like impact assessment (see Organisation for Economic Co-Operation and Development 2009 for a more extensive discussion).

- **Output** refers to the activities carried out in a given period. This dimension includes the absolute or relative (in per cent) number of primary laws and subordinate regulations appraised, the degree of compliance with guidelines, and the magnitude of analyses. Officers rely on various guideline documents for IA, specific tools used in IA such as benefit-cost analysis, and standards for consultation. Of course, guidance on new regulatory tools can be of good or bad quality (Radaelli and De Francesco 2007), which is why government bodies in the UK even produce guidelines on how to design better guidance (Better Regulation Executive 2009).

- When we think of **intermediary outcomes** we are concerned with behavioural and cognitive change, considering how regulators and inspectors perceive regulation as well as how citizens and the business community rate policy appraisal and, more generally, the regulatory efforts of the government. Data on inspections and enforcements also belong to this category.

- Finally, the causal chain should also show the effects of regulatory reform on **final outcomes**, such as number of new firms created and other classic economic
indicators that are causally linked to regulatory activity. Thus, if a government is committed to simple, targeted, efficient, proportionate and fair regulation, we would expect to find causal effects downstream.

Consequently, a high quality IA is based on robust and consistent standards, and it is grounded in robust economic analysis; but we should also see the results of the assessment activity in the final regulations. However, we can have good quality IAs and regulations that are not efficient, simple, proportionate, and fair. This may be the result of different variables at work. For instance, we can reason that IA is produced according to the guidelines, that measurement takes place, that the estimations of costs and benefits are overall accurate and realistic. And yet, regulators may not actually use the IA when taking decisions. Organisations can be ‘hypocritical’; they can produce evidence that is not used in their decision-making process (Brunsson 1989). Regulatory agencies may be under pressure to produce evidence, but may then use political thinking to steer the course of regulation. Instead of evidence: thus, we end up with politics-inspired policy instead of evidence-based. Even independent regulatory oversight bodies may adopt political criteria when they review the IAs produced by agencies, as some authors working on the US case have argued in the past (most notably West 2005 on the internalisation of political preferences in the US Office of Information and Regulatory Affairs). These bodies may delay regulation or ask agencies for more analysis because the proposed regulations are not in line with the regulatory philosophy of the White House.

To simplify these complex relationships, we discuss two dimensions of IA. We have two independent variables along the X and Y axis that produce an outcome in the four cells (our dependent variable). This is a common typological device called ‘explanatory typology’ (Elman 2005).

The first independent variable concerns the IA document. Does the impact statement comply with guidelines on policy appraisal? Do the estimates of costs and benefits reflect the best available evidence? Are tools such as benefit-cost analysis used correctly? Has regulatory bias been reduced in the final choice of an option or another?

The second refers to the IA usage in decision making and regulatory oversight, i.e. whether policy makers actually use the IA or leave it to gather dust in a drawer. We distinguish two types of usage here: to inform political decisions (evidence-based policy making) and to control the bureaucracy (regulatory oversight).

With regards to evidence-based policy making, authors studied whether regulators use the evidence contained in the IA and how the latter informs the legislative or regulatory decision-making process (Carroll 2010; Coletti and Radaelli 2013). Of course, ‘using
evidence’ is a rather ambiguous term. As the literature on knowledge utilisation (Boswell 2008; Radaelli 2010a; Schrefler 2010; Dunlop et al. 2012; Weiss 1979) suggests there is an instrumental type of utilisation when regulators draw on IAs to improve the contents of the regulatory choice, and a more strategic usage where evidence from appraisal is used to legitimate a pre-existing political preference. We know that policy-makers often learn politically: They use information and evidence to tweak policy with the expectation to increase popularity and win elections, not with the aim of improving on the efficiency and fairness of policy (May 1992).

• As to regulatory oversight, scholars explored whether governments use IA, and key tools in policy appraisal such as benefit-cost analysis, to exercise control of the bureaucracy (McCubbins, Noll and Weingast 1987; Horn and Shepsle 1989; Posner 2001; Shapiro 2005). More generally, it is plausible to assume that RIA is a tool to increase the power of what has been variously defined as ‘the Presidential administration’ in the US (Kagan 2001) or the ‘core executive’ in the UK (Dodds 2006). Accordingly, the main political effect is an increase in the power of the central structures within the executive.

By combining these two dimensions – the quality of analysis and usage – we generate four cases (see Figure 1).

![Figure 1: Four-Fold Typology of Impact Assessment Types](image)

‘Perfunctory’ IAs are a ritualistic exercise (quadrant 1): there may be provisions for IA but actually very little analytic work is carried out in the preparation of regulation, and evidence is not systematically reported in documents that can be discussed by cabinet committees or regulators (a phenomenon reported by Radaelli 2009).

Next to this are the cases are when bad analysis is used to inform decisions. Imagine regulators that draw on ‘back-of-the-envelope’ rough estimates’ (quadrant 2). Most likely, they do this because
they know that the low quality evidence has been produced to support their pre-fabricated preferences about regulatory choice. In these politicised circumstances, it doesn’t matter whether the analysis has depth and robustness, provided that there is some analysis that goes in the direction favoured by the regulators. This is of course a source of high political bias in political choice. It opens the door to all sorts of manipulation of what is supposed to be an evidence-based process.

When there is robust analysis but the regulators do not look at the evidence, the IA is reduced to little more than ‘background homework’ (quadrant 3). In this case, policy analysis does not speak the truth to power (to paraphrase Wildavsky 1979) either because the policy-makers have no faith in regulatory policy appraisal or because the evidence runs counter to the political priorities of the decision-makers. There are also pragmatic reasons why this can happen, not necessarily linked to political preferences. It can be that within an agency (McGarity 1991) or inside a government department there are different types of professionals. In some European countries economists write the IA, but the development of policy proposals is territory for the lawyers. Often there are different bureaus, one that prepares the IA and one that develops regulation. When this happens, even good IAs may end up being considered like background studies, but the real decisions are taken by different officers using different types of arguments and reasons.

The last cell for consideration (quadrant 4) is the optimistic scenario where IAs are robust and inform the regulatory decision-making process. We called this a case of ‘policy learning’ because regulation is managed with the tools that are supposed to provide the evidence-based input to decision-makers. The only problem here may be one of time alignment: the decision may be urgent, and the analysis may take time. We know that the time lines of knowledge production and policy decisions rarely run on parallel lines (see for example, Dunlop 2010 on the policy appraisal of biofuels). So, for example, if life-savings regulations are needed in a short period of time, it may be unwise to wait for the long time necessary to sift evidence and to peer review economic analysis. This leads to the principle of proportionate analysis as understood in IA: the depth of analysis must match the characteristics of the problem under consideration. Analysis ought to be as good as the entity of the regulatory problem suggests.

3. Reviewing the literature

We review the literature using our two dimensions of analysis and utilisation, noting that it is not common to combine the two. To be clear, what follows is not a comprehensive review, but rather we have selected key studies that exemplify the different approaches to exploring the quality of impact assessment.
Quality of analysis

Policy appraisals cover a variety of policy areas, many of them highly technical in nature. In order to replicate IA analyses, and thereby assess their quality, scholars would require both disciplinary and interdisciplinary knowledge and skills in areas as diverse as animal health, radio waves or corporate taxes. Although some IA scholars engage in in-depth, qualitative analyses of specific policy appraisals, the majority tend to use proxies for quality of analysis. We distinguish three types of proxies: the presence of guideline items (e.g., problem definition, estimation of costs, etc) in an existing IA, measured with scorecards; questionnaires exploring features of the IA deemed relevant by the scholar, and, finally, qualitative information is then converted into quantitative data.

Scorecards are arguably the most widely used approach to study quality. These benchmarks concern the analysis of samples of IA to establish whether they provide the information that they are supposed to convey to decision makers, according to IA drafting guidance documents and good international practice. An important caveat is that by examining the quality of IA in this way one cannot say anything about the plausibility of the analyses presented. After all, regulators may tick all the boxes required by the guidelines but deliver a poor analysis based on weak or flawed data. On the other hand, scorecards may be attractive to non-specialists in highly technical policy fields. Plus, the scorecard approach is not very time-consuming, enabling scholars to evaluate a large number of IAs and thereby improving the validity of the results.

Hahn and collaborators were the first to use scorecard approaches in their study of US IAs. In line with US official guidance, their scorecard puts a high premium on economic impacts, benefit-cost analysis, point and range estimates and cost effectiveness. Hahn and Dudley (2004) look at 55 IAs prepared during the Reagan, George H.W. Bush and Clinton administrations, reporting slight improvement over time although average scores were often low. Hahn (2005), referring to earlier work, discusses an evaluation of 168 IAs prepared between 1981 and 1986; further studies include Hahn and Tetlock (2008). Hahn’s work was highly influential to the study of policy appraisal but appears limited. It reflects the somewhat narrow perspective of US IA which mainly focuses on economic impacts of new regulations and their quantification and monetisation. In contrast to, say, IAs prepared by the European Commission, US IAs – and therefore Hahn and collaborators – say little about impacts on the environment or society more broadly. Clearly inspired by Hahn’s work, Fraas and Lutter (2013) assessed an evaluation of IAs prepared by the US Environmental Protection Agency.
Renda (2006) used the scorecard approach in an EU context, studying all 70 IAs produced by the European Commission between 2003 and 2005, the first years after the EU adopted this approach to policy formulation. The scorecard consisted of more than 200 items, documenting the breadth and scope of policy appraisal at EU level. Apart from cost-benefit analysis, Renda’s scorecard included the assessment of economic, environmental and social impacts, the presence of tests such as risk analysis or the precautionary principle, the analysis of alternative policy options but also EU-specific features such as the common market or the *acquis communautaire*. The European Commission’s early IAs – Renda concludes – were relatively poor with regards to the analysis of costs and benefits, rarely considered alternative policy options, did not provide much information on administrative burdens and struggled with recommended tests such as sensitivity analysis. More recently, Renda (2011) updated his data and argues that European Commission IA’s have improved significantly between 2003 and 2009. The assessment of economic, environmental and social impacts has become more systematic and complete with IAs quantifying and monetising costs and benefits more frequently. Although the 2006 and the 2011 study come with reflective sections on the EU policy context, the overall ambition of these works is descriptive rather than explanatory (Dunlop 2006). This also implies that quality of analysis is not discussed in relation to usage. Torriti (2007) offers an alternative study on early European Commission IAs, by and large confirming Renda’s findings.

Cecot et al. (2008) compare the findings of Hahn in the US and Renda in the EU. In a similar fashion, Fritsch et al. (2013) compare European Commission IA’s, scored by Renda and collaborators, with a unique dataset of 477 IAs prepared by UK government departments and agencies between 2005 and 2010. This is the largest dataset so far used for the evaluation of IA. Using a scorecard of 90 guideline items, the comparison suggests that policy appraisal in the two jurisdictions has improved steadily across the years. UK and EU IAs show similar patterns across a number of dimensions, for instance when it comes to the analysis of economic impacts and the identification of costs and benefits. Staronová (2010) presents a comparative study on IA in the Czech Republic, Estonia, Hungary, Slovakia and Slovenia.

To summarise, scorecard approaches usually share three characteristics. First, they are descriptive, i.e. they neither explain differences in quality of IA nor do they take quality of analysis as an independent variable and explore the implications of good or bad analysis for policy making. Second, scorecards focus on the quality of analysis; they do not study usage of IAs – which is why it is difficult to relate such work to our typology (introduced in Section 2). Third, scorecard approaches rely on a medium or large number of cases. Not surprisingly, datasets therefore contain
both excellent and poor IAs, making it difficult to generalise and, once again, to place a jurisdiction conveniently in one particularly quadrant of our typology.

The second proxy used to explore quality is the questionnaire – here we hear the voice of the scholar rather than official guidelines. At the UK’s University of East Anglia, have studied the quality of analyses in environmental policy appraisal in various European countries (Russel and Jordan 2007; Turnpenny et al. 2008; Russel and Jordan 2009; Russel and Turnpenny 2009). Their method builds on previous scorecard approaches when it comes to evaluating quality of analysis. However, the authors also carried out research interviews to explain varying degrees of quality across countries, thereby exploring causal question about IA. Russel, Turnpenny and collaborators identify, amongst others, organisational culture and prevalent policy styles as key obstacles to effective policy appraisal. Usage of IA, however, does not play a major role in their research, though (except Hertin et al. 2009 which does not talk about quality of analysis).

More recently, two US teams have departed from previous scorecard approaches. This scholarship uses questionnaires exploring features of the IA deemed relevant by the scholar rather than guideline requirements. Information gathered would then be scored on an x-point scale (the authors use various scales) and convert this information later into quantitative data for comparative purposes. What is more, the authors evaluate the quality of analysis mainly with a view to relate quality of analysis to other factors, i.e. to explore a causal relationship. Once again, usage of IA does play a role in those studies. Due to the large number of IA’s assessed, this research cannot find a clearly identifiable home in our typology; the IA’s assessed are simply too diverse.

- Shapiro and Morall III (2012) evaluate the quality of 109 US IAs published between 2000 and 2009 to correlate quality of IA analysis and the amount of net benefits identified in each IA, testing the hypothesis that the quality of IA analysis reflects the economic significance of a regulatory proposal. Work published in 2013 explores whether the time spent on regulatory oversight depends on the quality of analysis (Shapiro and Morall 2013). To this end, the authors use nine guiding questions, initially developed by the US Office of Management and Budget (OMB) – the national oversight body. The questionnaire covers, amongst other things, problem definition, policy alternatives, quantification and monetisation, uncertainty, and distributional effects. Interestingly, the authors found “a negative relationship between information provided by the analysis and the net benefits of the rule … [T]hose rules that most barely clear the net benefit threshold had the least useful analyses supporting them.” (p197).

- Based at George Mason University, Washington, a team directed by Jerry Ellig uses a questionnaire of up to twelve items to evaluate the quality of analyses in US IA. The questionnaire...
is similar to the one adopted by Shapiro and Morrall but includes a couple of additional questions, for instance whether citizens and other users would be able to locate, understand and verify assumptions, data and arguments. Ellig et al. (2012; 2013) also try to learn something about usage of IA, asking to what extent the analysis would affect decisions in the proposed rule. Because the authors study draft rather than final IAs, however, the answer is quite speculative; there is not empirical data on actual usage. Ellig and colleagues findings do not significantly change the overall picture presented by Hahn a couple of years before. However, the Washington team brings in a couple of nuances, for instances the difficulties experiences by US regulators to report adequately on the policy problem to be tackled or market failures identified. Key publications arising from this research are Belcore and Ellig (2008) on homeland security IAs, Ellig and McLaughlin (2012) on all IAs published in 2008, and Ellig, McLaughlin and Morrall III (2013) on all economically significant regulations between 2008 and 2010. In their later work the authors also try to isolate the effect of midnight regulations and ideological differences between the government and independent regulatory agencies.

Finally, let us have a brief look at qualitative case studies on IA quality. Generally, case studies trying to analyse, and perhaps replicate, one or two specific policy appraisals are still in great demand. However, see above, we have to keep in mind the difficulties related to this method. Examples include the work done by Harrington, Heinzerling and Morgenstern (2009) on three rules prepared by the US Environmental Protection Agency (EPA), and Torriti (2010) who studied the European Commission IA on the liberalisation of the EU energy market. It appears the analyses were weak in the studied cases, but usage is not discussed – hence we do not know whether those cases belong in quadrants 1 or 2 of our typology.

**Utilisation**

There are not many links between the literatures on the robustness and scope of analysis and usage. First, a few people only work on both dimensions of quality. Second, while both US and European scholars have contributed to the literature on analysis, work on usage is a bit more Europe-centric. This has implications for theory and research directions: a majority of scholars looks into evidence-based policy making and learning, but neglects the bureaucratic control and oversight dimension – the original contributions to this field oversight come from the US. Third, students of usage prefer to rely on qualitative case study designs. These approaches are difficult to reconcile with the strong quantitative focus of the quality of IA analysis camp.
There are plenty of suggestions about how IA can be used in policy making (see, for instance, Radaelli 2010a; Dunlop et al. 2012). We discuss two of them: evidence-based policy making (learning) and oversight (control).

With regards to learning, a simple assumption that follows from our typology, would be that IA is either used for learning – or not. Many authors do indeed emphasise the absence of learning. IA then take a mere symbolic role in the policy process (Hertin et al. 2009; Radaelli 2010a; Schrefler 2010; Rissi and Sager 2012). On the other hand, authors describe and theorise quite a few occasions in which IA contributed to learning in decision-making. Authors thereby distinguish three types of learning: instrumental, conceptual, and political. This distinction reflects earlier research on learning (see, for instance May 1992).

- As the ideal-type of evidence-based policy, instrumental learning occurs when IAs support decision makers through data, models, predictions. Authors provide plenty of evidence that this actually occurred in case studied (Hertin et al. 2009; Radaelli 2010a; Schrefler 2010).

- Conceptual learning describes the possibility that actors change, due to the IA, more fundamentally their way of thinking about specific policy problems and solutions. The literature reports much less instances of conceptual than of instrumental learning (Hertin et al. 2009; Rissi and Sager 2012).

- Political learning is a more strategic approach to using evidence generated in policy appraisals. Hertin et al. (2009), for instance, report of policy makers justifying pre-existing policy priorities through knowledge gathered in IAs. Other authors argue that decision makers represent knowledge in a one-sided way in order to strengthen their line of reasoning in public or use IA information to improve their position in the political realm more generally (Rissi and Sager 2012; Schrefler 2010).

- The question that we take from this literature therefore is not whether usage of IA occurs in evidence-based policy making. The question is: which kind of learning – and is this (partly) dependent on the quality of analysis, i.e. whether the evidence base is good or bad, disputed or widely shared. This leads us to consider causal factors influencing the type of learning that authors observed.

In a nutshell, the literature appears to be much more diverse if we look into causal factors proposed in the literature. Schrefler (2010) suggests, on the basis of three in-depth case studies carried out inside the UK telecoms regulator, that the level of policy conflict, public and parliamentary support for policy proposals, and the magnitude of and (non-)controversy about data, methods, models available in a policy field play a key role. We find the third factor – data in the widest sense –
particularly useful in this context, as this relates to the above question how quality of IA analysis and usage in policy making relate to each other. Nevertheless, the evidence is inconclusive so far, much more research is needed to refine this potential relationship (and we hope our typology may help here). Furthermore, Hertin et al. (2009) identified pre-existing political commitments, legal requirements and the interests of key actors as key drivers for usage of IA evidence in policy making (for a similar argument, see Turnpenny et al. 2008). Strong institutionalised commitments towards cost-benefit analysis and requirements to express policy options in economic terms would, for instance, discourage regulators to think outside the box and engage in conceptual learning. Sager and Rissi (2011), finally, emphasise the importance of different modes of democratic organisation. For example, there is generally less scope for expert-based input in direct-democratic systems such as Switzerland. To summarise, while exciting work has been done on usage of IA from an evidence-based policy perspective, these studies rarely conceptualise systematically quality of analysis as an influential factor. Consequently, it is too early to place previous scholarship conveniently in specific quadrants of our typology.

We found less scholarship dedicated to the usage of IA from an oversight perspective. We know that academics and policy makers alike emphasise the opportunity to use IA this way, we also know a lot about the respective oversight bodies in the US (West 2005), the EU (Wiener and Alemanno 2010) and the UK (Gibbons and Parker 2012, 2013). However, we know little about their effectiveness in controlling the bureaucracy and overseeing regulatory proposals. Previous work includes, on the one hand, Shapiro and Morrall (2013) who studied the relationship between quality of policy appraisal and time spent on drafting and reviewing IAs. They confirm a positive correlation exists but the causal direction is unclear to date. Review bodies might take longer to scrutinise long and rich IAs. Alternatively, however, IAs reviewed in a long process might, because of the long process, just become better than IAs that have received less attention. On the other hand, Radaelli (2010a, 2010b) explored the usage of IA in seven countries. Likewise, the evidence is inconclusive and “does not enable us to separate neatly the three ideal types of control, rationality and public management reforms” (p180).

To conclude, although there is a rich literature on both quality of IA analysis and quality (and types) of usage, there is little dialogue between these two strands of research. Our typology, introduced in Section 2, is designed to encourage research trying to tackle exactly this relationship.

4. Conclusions

In this article we have discussed the quality of IA and reviewed the literature available. We focused on two dimensions: the analysis and overall evidence-base provided by the IA, and what regulators
do or don’t do with the results of appraisal. By considering the two dimensions of analysis and utilization independent, we have generated four scenarios that help us focus the discussion of the literature.

One result of our reading of the literature is that there is a strong division of labour between those who work on one dimension and the other. This also reverberates on the choice of methods. It is easier to talk about usage of IA by using qualitative methods and case studies, whilst it’s natural to compile scorecards to generate data on the analysis contained in the appraisals produced by a given government.

The second result of our typological suggestion is that the next step is to link the two dimensions – but this has been done rarely up until now. Indeed, we can formulate testable hypotheses by linking analysis to usage. We argue that low quality analysis is more prone to manipulation and political learning in usage. Low quality analysis can also be used symbolically, for example to report to the OECD that ‘our country produced more than 100 appraisals of new legislation last year’. In short, there are many ways to utilization of low quality IAs, including getting legitimacy in international circles. And of course low quality analysis is easy to ignore, or forget in a draw at the office. Our hypothesis is therefore that low quality analysis tends to associated either with non-usage of manipulations. Robust analysis is more difficult to distort, hence the hypothesis in this case is that elected policy makers either ignore it altogether or use it to learn instrumentally.

The third result is that analysis and utilization are also useful for the discussion of a dilemma as old as the Weberian analysis of bureaucracy: do politicians control the bureaucracy? Can the bureaucracy escape the controls and exercise autonomy? With high quality analysis and coherent utilization the bureaucracy has an important role to play in regulatory choice. If the analysis is strong but ignored at the moment of choice, one could reason that IA is nothing more than a fire-alarm system to alert the constituencies supporting the decision maker. Unless the IA rings bells that alert the minister, there is no need to pull the alarm. The politician may also use IA requirements to saddle the agencies and delay regulation, or simply increase the cost in time and analysis (for the agency) to produce regulation: IA as grit that makes the regulatory machine more difficult to operate, and slower.

References


