THE CONTENT OF IMPACT ASSESSMENT IN THE UNITED KINGDOM: EXPLORING LEARNING ACROSS TIME, SECTORS AND DEPARTMENTS

Oliver Fritsch (University of Leeds, o.fritsch@leeds.ac.uk)

Jonathan C. Kamkhaji (University of Exeter)

Claudio M. Radaelli (University of Exeter)

ABSTRACT

Whilst several studies have documented how evidence-based policy instruments affect public policy, we know less about what causes changes over time in the analyses mandated by the instruments, especially in Britain. Thus, we take the analytical content of a pivotal regulatory reform instrument (impact assessment) as dependent variable, we draw on learning as conceptual framework, and we explain the dynamics of learning processes across departments, policy sectors, and time. Empirically, our study draws on sample of 517 impact assessments produced in Britain (2005-2011). Experience and capacity in different departments matter in learning processes. Guidelines matter too, but moderately. Departments specialize in their core policy sectors when performing regulatory analysis, but some have greater analytical capacity overall. Peripheral departments invest more in impact assessment than core executive departments. The presence of a regulatory oversight body enhances the learning process. Elections have different effects, depending on the context in which they are contested. These findings contribute to the literature on regulation, policy learning and policy instruments.

KEY WORDS

Regulation, Regulatory Impact Assessment, Evaluation, Evidence-based policy-making, United Kingdom

ACKNOWLEDGEMENTS

This paper is based on research carried out with the support of the European Research Council (ERC) project on *Analysis of Learning in Regulatory Governance* (ALREG, Advanced Grant no. 230267), directed by Claudio. The authors gratefully acknowledge the support of the ERC. We thank Jan-Pieter Beetz, Sarah Cooper, Samuele Dossi, Theofanis Exadaktylos, Anaid Flesken, Mattia Gallotti, Will Hankey, Stuart Ingham, Martin Moorby, Gabriel Thebolt, Siim Trumm and Sophia Zeschitz for research assistance.

INTRODUCTION

In the vast literature on policy instruments, Impact Assessment (IA) has a special place as a pivotal instrument in regulatory reform (Dunlop and Radaelli 2016a). IA is a tool to appraise the effects of proposed primary and/or secondary legislation. It has been adopted over the last 35 years by many OECD and EU member states (Radaelli 2005; OECD 2009; De Francesco 2012) and, increasingly so, developing countries (Adelle et al. 2016).

For social scientists, this regulatory policy instrument features prominently on two distinct, but not necessarily contradictory, political agendas: it is, on the one hand, a manifestation of the evidence-based policy making movement (Nutley, Walter and Daviers 2007; Cairney 2015); or, if we want to take the long view, an episode in the struggle to bring rationality and science to bear on public policy (Carley 1980; McGarity 1991; Boswell 2008; Carroll 2010; Desmarais and Hird 2014). On the other hand, it is about controlling the bureaucracy, as shown by the literature on administrative procedure (McCubbins, Noll and Weingast 1987; Carpenter and Gubb 2014; Damonte, Dunlop and Radaelli 2014). In both strands of the literature, IA plays the role of an independent variable with causal effects in terms of political control and knowledge utilization.

We take a different perspective and start from the content of IA as a dependent variable. In doing so, we contribute to an emerging literature that has measured the content of IA and addressed the question of whether regulators comply with statutory guidelines and international best practice (Cecot et al. 2008; Staronová 2010; Shapiro and Morrall 2012; Ellig, McLaughlin and Morrall 2013).

We analyze a sample of 517 IAs produced in the UK between 2005 and 2011. Our contribution to the comparative literature on policy instruments is *substantive* – because we explain variation across time, sectors and departments; *conceptual* – because we draw on a theoretical lens on the policy process, that is, learning; and *methodological* – because we provide a template of how a large-N dataset of IAs can be assembled, coded, and analyzed.

We proceed as follows. First we briefly introduce the literature and, in another section, the UK context. This is followed by a description of our dataset, coding technique and dependent variable. We then discuss theory, starting from a simple proposition about time. Further, we add a level of sophistication, by testing whether (a) the basic trend is mitigated or interrupted by statutory guidelines, regulatory oversight, and election years; (b) departments intensify regulatory analysis of costs and benefits affecting the policy sectors they specialize in; and (c) IA is sensitive to whether a department belongs to the core executive or not. We finally discuss our results and their contribution to the literature.

LITERATURE

The studies on the analytics of IA have often stop at simple compliance tests: do regulators comply with the statutory guidelines on regulatory analysis or not, could benefit-cost calculations be carried out more effectively by agencies (for the US, see Hahn and Tetlock 2008; Belcore and Ellig 2008; Ellig and McLaughlin 2012; Fraas and Lutter 2013; and for the European Commission, see Cecot et al. 2008 and Renda 2011). The approach features also in Fritsch et al. (2013) on the UK and Staronová (2010) on four central and eastern European countries. Others have looked at the content of IAs from a different angle, that is, whether they assist in meeting the goal of mainstreaming some key dimensions like gender, fundamental rights, and so on (on the EU, see Smismans and Minto 2016). These studies, albeit descriptive, "provide a starting point for analyzing a variety of factors that might influence the quality of regulatory analysis, such as the nature of the regulation, politics, legislative mandates, or deadlines" (Ellig and McLaughlin 2012, 863).

Essentially, in the literature these factors are examined from three angles. First, the diffusion perspective (De Francesco 2012; Wiener 2007; Peci and Sobral 2011) sheds light on the interaction between international organizations and adopting countries. We do not borrow from this literature. The UK is a pioneer country in regulatory analysis. Together with the US, it is an exporter, not an importer of IA.

Second, quantitative work carried out in the US focuses on single causes rather than a range of hypotheses. For instance, Shapiro and Morrall (2012) examine 109 IAs to test the hypothesis that the quality of analysis reflects the economic significance of a regulatory proposal. Analysing 111 IAs, Ellig, McLaughlin and Morrall (2013) isolate the effects of midnight regulations and ideological differences between government and regulators. Shapiro and Morrall (2013) explore the relationship between time spent on regulatory oversight and quality of analysis.

Third, the literature on environmental impact assessment is another source of inspiration (e.g., Tzoumis and Finegold 2000; Gray and Edward-Jones 2003; Tzoumis 2007; Pinho, Rodrigo and Monterrosso 2007; Landim and Sánchez 2012; Kabir and Momtaz 2014). Providing evidence on countries as diverse as Bangladesh, Brazil, Portugal, the UK and the US, this scholarship provides in-depth analyses of trends over time and across departments, and identifies variables that explain content, such as legislation, guidance documents, agency age and experience, staff turnover, and time spent on appraisal. Studies of individual jurisdictions dominate (but see Canelas et al. 2005 for a comparative project).

We draw three preliminary conclusions. On the one hand, the content of IA has been studied in various contexts, but this strand of research is often descriptive or associated with one or two variables only. On the other, IA studies usually rely on rather small datasets or qualitative research. Finally, research on environmental impact assessment has identified a number of causal factors that should be tested beyond the field of environmental and sustainability appraisal. We proceed from these findings and try to overcome the limitations in the literature by creating a new dataset and by properly testing six factors that may affect the content of regulatory analysis.

REGULATORY ANALYSIS IN THE UK

IA is a mandatory step in the UK policy formulation process, covering both primary and secondary legislation (of regulatory nature or not; hence we use IA rather than RIA). While the bulk of analyses performed within IA since the 1980s focused mainly on administrative burdens for businesses, more recent statutory guidelines have expanded the scope of IA, thereby including further analyses relating to competitiveness and the position of small and medium enterprises, environment and sustainability tests, and tests on public health, social welfare and vulnerable societal groups (Cabinet Office 2003; Department for Business Enterprise and Regulatory Reform 2007). In recent years, the production and scrutiny of IA has been key to deliver on the government's major regulatory policy target: to reduce the impact of regulation on business. This was done, first, through 'one-in-one-out', an initiative which required departments to accompany proposed new regulations with deregulatory measures of the same net cost (HM Government 2011). 'One-in-two-out', adopted in 2015, takes this objective even further, obliging regulators to offset every pound of cost imposed through new regulation with deregulatory measures worth twice as much (Department for Business Innovation and Skills 2015; Lodge and Wegrich 2015). IAs then serve to find out if new regulations comply with the 'equivalent annual net cost to business' figure used for the 'one-in-two-out target'.

The UK has developed one of the most advanced IA systems in Europe (Hertin et al. 2009; Renda 2011). Since 2009, the Regulatory Policy Committee (RPC), an advisory non-departmental public body, has scrutinized new regulatory and legislative proposals at the drafting stage and, if necessary, suggests modifications, additional tests or new data to IA officers (see Regulatory Policy Committee 2010). Its mission has widened to connect IA with government targets like 'one-in-one-out' and recently 'one-in-two-out' (by checking claims made by departments about how new proposals met the one-in-one-out and, today, one-in-two out requirement). The RPC also features prominently on the government agenda because it validates the governmental estimates of costs and benefits arising out of new regulations (Gibbons and Parker 2012, 2013).

Turning to empirical research on the UK, Ambler, Chittenden and Xiao (2007) report on categories of costs and benefits, arguing that incomplete or inaccurate analysis result in heavier regulatory burdens for business. Fritsch et al. (2013) compare UK and EU IAs, concluding that the EU outperforms the UK when it comes to less orthodox tests in IA, such as environmental and social impacts.

When we think of IA as a regulatory tool more broadly, the literature has for a long time centred on political control over regulators. Historically, this theme has revolved around the relationship between the Cabinet Office, Whitehall and the regulators. Both under Labour and in the Coalition government, Britain has witnessed an increase in the control capacity of the Cabinet Office (Dommett and Flinders 2015), seeking to reduce fragmentation and re-establishing control at the centre through the obligation to carry out reviews, evaluations and IAs of new legislation and regulation. Indeed, the unmitigated faith in IA goes back to the 1990s, when a study on compliance cost assessment (Froud et al. 1998) demonstrated how the core executive tried to achieve its objective of de-regulation by exercising regulatory oversight via analytical tests on new regulations.

There is of course a wealth of studies on regulation in the UK (Ogus 2001; Lodge and Wegrich 2009; Prosser 2010), yet we know little about the crucial moment in which regulation is appraised via IA (but see Russel and Jordan 2009 on IAs as levers in policy coordination for sustainable development).

DATA, CODING AND DEPENDENT VARIABLE

This section describes our dataset, coding technique and our dependent variables. We follow the OECD convention of considering analytical richness, which refers to breadth and scope as key dimensions to differentiate IAs in terms of the information they provide (OECD 2009). Scope refers to the number of elements (e.g. problem definition, economic, environmental and social analysis, consultation) and breadth to how deep the analysis goes (e.g. benefit-cost identification, quantification, monetization, benefit-cost ratios, discount rates and sensitivity analysis).

As mentioned, we built a sample of 517 IAs between 2005 and 2011. First, we established a database of all IAs produced by the UK central government in this period, leaving aside Scottish IAs which belong to a different system of regulatory oversight in Edinburgh. For 2005 to 2008 we perused government command papers on regulatory analysis, i.e. documents sent by the Cabinet Office to Parliament to inform about on-going legislative or regulatory activities. For IAs produced between 2008 and 2011, we used the Impact Assessment Library published by the UK Better Regulation Executive. In total, the database lists more than 2,000 entries. In a second step, we extracted a stratified random sample of 517 IAs, representative across departments and time and proportionate to the productivity of departments in a given year. We then downloaded all IAs in our sample from departmental websites and the IA library or approached government departments via email in order to get hold of those documents.[i]

We relied on the scorecard approach (Hahn and Tetlock 2008; Shapiro and Morrall 2012; Ellig, McLaughlin and Morrall 2013) to appraise whether IAs provide the information that they are supposed to convey. Taking government guidelines as a benchmark, we established a list of typical tests and analyses. The list includes varieties of economic tests, policy effects on the economy, health, labor and the environment, but also the definition of the policy problem, consultation, and issues related to implementation and monitoring. In total, our scorecard covers 79 items.[ii] Trained graduate coders piloted the system under the authors' supervision, with checks on their reliability scores and additional advice. After this phase, the coders carried out the measurement in binary scale (0/1, absence or presence) for all scorecard items. We then measured again the reliability of coding and variance across the coders throughout the coding stage of our project.

Note that the fact that a department reports in a given IA that a cost quantification has been carried out is all a researcher can code. Our coders could not answer the question whether the analysis rested on sound assumptions, solid figures and appropriate statistical techniques. In the end we can only check for compliance with the requirements set by the IA guidelines and international best practice.

The previous discussion of the literature suggests that analytical richness is potentially influenced by different factors. We concentrate on the following six: experience, learning through guidelines, regulatory oversight, sectoral specialization, elections, and political priorities of the core executive. In order to test these causal claims, we created two indexes that work as dependent variables:

one that aggregates the information on all 79 guideline requirements and represents a proxy of the overall analytical richness; and one set of four sub-indexes which focus on specific guideline requirements and represent proxies of, respectively, completeness of benefit-cost measures, and economic, environmental and social analysis. We construct our main index in two steps:

First, we perform a Principal Component Analysis (PCA) to minimize redundant information, i.e. to reduce the number of manifest variables to a smaller set of components characterized by a simple structure and explaining a satisfactory degree of variability. The PCA results in 15 significant components explaining more than the 50 per cent of the overall variability observed in the sample. Second, we use the 15 principal components and their scores to construct a weighted index (for instance, Vyas and Kumaranayake 2006) which reads as follows:

$$Weighted\ Index\ Score_i = \sum_{j=1}^{N}(component\ score_{ij}\ \cdot\ variance\ explained_j)$$

whereby the subscript *i* indicates the observation and the subscript *j* the principal component (N=15). The index is a hierarchically weighted aggregate of the 15 components. Because we are interested in variation, the components explaining major shares of variance in the data set carry a greater weight. Finally, we rescale the new scores on a 0-1 scale to make interpretation easier.[iii]

We follow a similar rationale to construct the sub-indexes. Each sub-index measures one dimension of regulatory analysis only. To illustrate, sub-index 1 on benefit-cost measures is based on those principal components that aggregate variables dealing with benefit-cost analysis, that is, components 1 and 2.[iv] Likewise, sub-index 2 on economic impacts aggregates only those principal components that cluster scorecard variables related to impacts on business, trade, GDP, and competitiveness (namely components 4, 7, 11 and 13). Finally, sub-indexes 3 and 4 on the analytical richness of social and environmental analyses brought together, respectively, the components related to social impacts (5, 8, 12 and 15) and those related to impacts on the natural environment (6 and 9).

In the following, we use those weighted indexes - proxies of analytical richness of IA in general and in specific dimensions of analysis - as dependent variables, enabling us to put the above hypotheses to an empirical test.

INDEPENDENT VARIABLES AND HYPOTHESES

We perform tests on variation across departments and over time by keeping the institutional context constant – one advantage of considering a single country. Essentially, we start with the baseline analysis of whether compliance with the requirements increases over time. We then consider the different variables that may alter or disrupt this process.

Hypotheses 1: learning through experience

To begin with, we draw on Kelman (2005). This author documented that change has a mundane, ordinary yet by no means trivial, characteristic: by simply doing ‘something new’, public managers show to other managers that a certain function or operation is feasible and can be carried out without too many impediments. Once we enter a new procedure in public administration and that procedure starts being used, experience plays an important role in the implementation process. We do not simply look at the passing of time (like in Fritsch et al. 2013). Instead, it is the accumulated experience in doing IA within a specific time frame that may be a cause for change. Alternatively, experience may be brought in externally through the appointment of new talent, and it may also be a result of intra-departmental specialization and selection, if authors of excellent IA are asked to work on other IAs in the future. We cannot possibly test those three mechanisms, not the least because we are likely to observe them at the same time. However, what all three mechanisms – learning by doing, new hires, intradepartmental selection – have in common is that they occur in response to ‘numbers’: officers learn more as they prepare more IAs. Departments are more likely to recruit new talent as the workload associated with IA preparation increases. Departments are more likely to allocate responsibilities to specialist authors within their institution as the number of IAs on their desks and need for special skills increases. This idea lies at the heart of hypothesis 1 which stipulates that the analytical richness of IA depends on the number of IAs prepared previously in a department:

H1: The more IAs a department produces, the better the department becomes in regulatory analysis.

Hypotheses 2 and 3: learning through experience – variations

H2 through H6 are in different ways variations in the mechanism underlying the first hypothesis, especially and most directly H2 and H3. Hence we deal with H2 and H3 together. Learning is facilitated, or hindered, by various political and organizational factors. We explore two of them. One is the introduction of new statutory guidelines on IA. More specific instructions should generate improvement in compliance (*Hypothesis 2*). Second, if the government creates a regulatory oversight body to check on the quality of IAs we could expect some effects. We reason that the establishment of the RPC in 2009 has enhanced regulatory analysis either through feedback and advice on draft IAs provided or through better IAs prepared in anticipation of RPC peer review (*Hypothesis 3*). We therefore hypothesize:

H2: After the release of statutory guidelines the IAs become analytically richer.

H3: IAs produced after the establishment of the oversight body, the RPC, are analytically richer than IAs produced before.

Hypothesis 4: election years

At the end of an administration, regulators are under pressure to get so-called midnight regulations out of the door (Beermann 2009; for an extension and test on IA, see Ellig, McLaughlin and Morrall 2013). This pressure, the argument goes, leads to poorer IAs when we get closer to the electoral deadline – hence the learning process is interrupted. We elaborate on this logic - and look at dawn as well as midnight. In the first months after elections – we submit – the bureaucracy waits for signals on the regulatory philosophy of the executive. This causes a sort of relaxation in the analytical richness of the IA. This argument is most likely flawed in the US, where Presidents typically issue executive orders on IAs in the first months, if not weeks, of their administration. However, there is nothing like that in the UK, so there probably is uncertainty after elections. We therefore also consider a modified version of our first hypothesis covering election years – which is characterized by two effects, midnight and dawn, in the same direction.

H4: Over time election years affect the trend in compliance with the requirements for IA.

Hypothesis 5: core issues and constituencies

The content of specific IA tests may depend on issues and constituencies. Environmental regulators – such as the Department for Environment, Food and Rural Affairs – may invest more in the analysis of the environmental impacts than, for instance, the Department for Work and Pensions or the Cabinet Office, especially in periods of austerity. Classic studies on the bureaucracy (Downs 1966; Wilson 1991) show that there is signification variation in how individual agencies and departments behave, but in the end they tend to develop routinized relationships with their external environment, especially the populations they regulate, serve, or control. Regulators build and cultivate their reputation in the constituencies they serve (Carpenter 2010). A department that has industry as key constituency will go deeper in the analysis of costs, especially costs for industry, much better than the Department for Work and Pensions, which is most likely to put a high premium on impacts related to social welfare. Environmental or health departments will take care of sustainability or public health much better than the median department. We are not in a position to say which of these factors feature most prominently. However, we can test the claim that regulators learn to carry out IA unevenly, thereby reflecting patterns of departmental specialization. This leads us to:

H5: Departments do not implement IA requirements evenly. Their analyses of the portion of the requirements that reflects a departmental core mission are analytically richer than their analyses of the portion of the requirements that does not relate to a departmental core mission.

Hypothesis 6: position of individual departments within the executive

We can also test whether the regulators are sensitive to the political priorities within the executive. Let us consider that core departments, tasked with designing broad government policy agendas and controlling public expenditure (think of the Cabinet Office or the Treasury), are more likely to support regulatory analysis. Classic regulatory departments like the Health and Safety Executive or the Department for Work and Pensions should be less enthusiastic about learning how to perform regulatory analysis. Typically ministers that feel strongly about controlling the budget and limiting expenditure are part of the core executive, whilst regulatory departments are somewhat peripheral to the key expenditure control agenda of the Prime Minister and the Treasury in the UK. We therefore hypothesize:

H6a: IAs carried out by departments close to the political centre of government are analytically richer than IAs carried out by peripheral departments.

And yet, even if the degree of government control over departments has an effect, the direction of that effect is far from being certain. One could argue that key departments are close to the political agenda of the Cabinet Office, and will not feel under pressure to perform good regulatory analysis. It is the peripheral department that is obliged to report on various effects of their policies before getting clearance by cabinet committees. The observable implication of this counter-argument is the following:

H6b: IAs carried out by peripheral departments are analytically richer than IAs carried out by core departments.

In the following, we put these propositions to an empirical test.

HYPOTHESIS 1: EXPERIENCE

Let us start with Figure 1 which visualizes the main index between 2005 and 2011, showing the richness of IA in this period. Although the effect is rather small (R2 linear =.035), Figure 1 suggests that, since 2005, IAs have become richer in analysis.

This finding is also supported by trend analysis. Based on a one-way ANOVA with a polynomial contrast up to the 5th grade and publication years of IA (in our sample: seven) as grouping variable, we identify a statistically significant positive linear trend of analytical richness over time ($p<.01$). We also observe a significant quartic trend at a 10 per cent confidence interval, indicating that the positive linear trend might be subject to bends and blips, highlighted too by the graph.[v]

--- FIGURE 1 ABOUT HERE ---

Intuitively, this could be interpreted in three ways. a) Learning through experience: over time officers become more familiar with the requirements for policy appraisal and develop their analytical capacity. They learn how to carry out tests, build capacity to obtain data and supportive materials, and consult better. b) Learning through specialization within departments: officers who have prepared excellent IAs in the past are more likely to be asked again to prepare another. c) Learning through new hires: in response to poor IAs prepared in the past, departments may decide to recruit more qualified individuals. To be clear, we are not in a position to test which of those three intuitions explains the analytical richness of IA in our sample best. This would require qualitative data that is difficult to obtain in a large-N setting.

Yet all three intuitions invite an important question: what happens 'in time' that makes officers learn how to prepare better IAs; why would departments select in-house specialists for policy appraisal or recruit new talent? Surely, officers do not learn how to use policy instruments by tearing off calendar pages. They learn by preparing analyses, by consulting the guidelines, by receiving internal and external feedback. The implication, then, becomes: the more IAs the departments produce, the more experience they gain, resulting in better IAs over time. Likewise, it is plausible to assume that intra-departmental processes of specialization depend on the number of IAs a department produces on average; and so does the willingness to recruit new, well-qualified staff. We therefore assume that a very productive department (say, 20 IAs per year) improves quickly whereas a less productive ministry (say, 4 IAs a year) requires more time to reach the same standard.

We operationalize this claim through departmental 'stacks'. Using the publication date of each IA, we sorted all impact statements in our sample, chronologically and by department. We then compared 'stacks' of IAs: the first 5 IAs of a department, the first 10, the first 20, the first 30 and so forth. The index, so we hypothesize, improves as a department appraises more and more rules. We selected eight departments, all of which in operation since 2005. Table 1 below summarizes our findings:

--- TABLE 1 ABOUT HERE ---

What does the data tell us? First, we compare the first and the last stack of each department and observe improvement in five out of eight cases, most notably for the Department for Environment, Food and Rural Affairs and the Food Standards

Agency. In other words, the more IAs those five departments completed, the better they became at providing information. In two cases, however, the Department of Transport and the Home Office, earlier IAs were better than later ones, and there is no significant development over time for the HMRC.

Second, although there is evidence that overall departments produce better IAs as they gain experience, this trend is not observed in all departments. The IAs of the Department for Trade and Industry, for instance, show an erratic pattern whereby weaker IAs follow better ones and *vice versa*. Third, the first five to ten IAs seem to be crucial: many departments made a step forward here but then reached a plateau without further development.

Finally, the IAs of some departments were already of high quality in the first ‘stack’, just see for example the IAs prepared by Home Office, whereas others took some time to reach the same standard. Apparently, departments do not start at a similar baseline and improve, whereby the degree of improvement then depends on the number of IAs prepared. Quite the contrary, while there is some evidence to suggest that the continuous production of IA contributes to the building of institutional capacity, we are reluctant to make a strong case: It seems the hypothesis can only be confirmed for departments starting at a lower baseline. Ministries that have already begun at a higher standard quickly reach a plateau characterized by no or little further improvement. This may be because they developed analytical capacity for appraising policy before the period under consideration here – e.g., a certain department may have historically invested more in capacity for economic assessment of proposals, independently of IA requirements (see Dunlop and Radaelli 2016b on different notions of capacity in the context of policy learning). No doubt other factors need to be discussed – as we do in the remainder of this article.

HYPOTHESIS 2: LEARNING THROUGH GUIDELINES

Thus, can we say that organizations learn? Perhaps. Sometimes. But the incremental development of in-house capacity – through learning-by-doing, intra-departmental specialization, or new appointments – is probably not the only and certainly not the most relevant factor. After all, if person A completed an HMRC IA in 2005, and person B prepared an HMRC IA in 2006, it is difficult to argue that there is *per se* a learning effect over time – unless there are mechanisms in place ensuring that previous experiences are passed on, repeated, refreshed. Unfortunately, we do not possess data on training events offered on IA. However, guidance documents are an alternative way of passing on knowledge to new generations of officers. One can suggest that, when the Department for Business Enterprise and Regulatory Reform released its 2007 guidelines, IAs have improved. We carry out two tests to explore this claim.

First, the 2007 guidelines included more precision on how to carry out cost-benefit analysis and a template with an overview page which summarized key findings on total costs and benefits and similar tests. The new summary page was supposed to remind officers of several important tests to be carried out before finalizing the IA.

Consequently, we only look at IA sections dealing with costs and benefits to see whether they have improved after the adoption of the 2007 guidelines. To this end, we use sub-index 1, aggregating data on measures related to the quantification or monetization of costs and benefits in IA, and create a dummy variable to contrast pre- and post-guideline IAs.

The results of the t-test, displayed in Table 2 below, indicate that respective IA sections became significantly richer after the introduction of the guidelines.

--- TABLE 2 ABOUT HERE ---

Second, we compare the overall quality of IA before and after the adoption of the 2007 guidelines, assuming their innovations informed analytical steps more generally. To this end, we use the same dummy variable employed before to contrast pre- and post-guidelines IAs. We hereby expect a cohort effect, according to which IAs published after the adoption of the new guideline are, on average, analytically richer than those completed before. Tables 3 below reports our findings.

--- TABLE 3 ABOUT HERE ---

We find that IAs pre-guideline are generally less rich than post-guideline ones. However, the effect size value is not impressive. This suggests that IA guidelines, at least in the period we examined, supported the learning process, but only moderately.

HYPOTHESIS 3: LEARNING THROUGH REGULATORY OVERSIGHT

Statutory guidelines are not the only mechanism of learning established by the government. The Regulatory Policy Committee (RPC) appraises draft IAs using five criteria: problem definition, presentation of options, evidence base, cost-benefit analysis, and overall presentation. One can therefore argue that the RPC has enhanced regulatory analysis, either directly through feedback and advice on draft IAs or indirectly, i.e. regulators go deeper and wider in their analysis in anticipation of RPC scrutiny.

In order to probe this intuition, we compare the values of the main index before and after the establishment of the RPC. We create a dummy to distinguish pre- and post-RPC establishment IAs and perform a t-test. Tables 4 below summarizes our findings.

--- TABLE 4 ABOUT HERE ---

They do indeed suggest that post-RPC IAs are richer in analysis than their pre-RPC counterparts – thereby supporting the intuition.

HYPOTHESIS 4: LEARNING IN ELECTION YEARS

The adoption of midnight regulations, so we hypothesize, may come with more superficial IAs. Likewise, in the first months after the election the new ministers will press hard to send signals to their constituencies with the swift adoption of new regulations, which will be supported by sub-standard analysis.

General Elections were held in the UK on 5 May 2005 and 6 May 2010. Trend analysis and comparison of estimated marginal means (see Table 6) has already suggested that, along a significant linear trend, there was a drop in the index in the election year 2010. We then performed a further ANOVA with planned contrasts to explore the effect of electoral years on IA. The tested contrasts capture the effects of the 2005 and of the 2010 general elections, see Tables 5 and 6 below (for a general discussion of this method, see Seltman 2015[vi]).

--- TABLE 5 ABOUT HERE ---

The effects, i.e. the differences between the mean values of the index in contrasted years, are statistically significant[vii] and indicate, in particular, that in 2010 the analytical richness of IAs has significantly decreased. In other words, public managers did indeed produce poorer IAs in election years (Table 6).

--- TABLE 6 ABOUT HERE ---

However, we do not know yet whether this is caused by midnight or dawn regulations. Let us study the timeline in more detail: Tables 7 and 8 below display the index in the months before and after the 2005 and 2010 general elections. Months highlighted in red are below the annual average, months highlighted in green are above.

--- TABLE 7 ABOUT HERE ---

According to our 2005 data, dawn regulations did not negatively affect IA quality. True, departments completed a handful of substandard IAs immediately after the elections in May. However, these were minor policy initiatives on milk pricing and land drainage improvements that can hardly be interpreted as pet projects of policy-makers. At the same time, we have some above-average IAs published four weeks before the elections in April 2005, and the January and February IAs are only slightly below the national average – suggesting that midnight regulations had only minor effects or no effect at all.

--- TABLE 8 ABOUT HERE ---

The 2010 data is more interesting. Again, there is no supporting evidence that midnight regulation made IAs poorer. IAs in late 2009 and early 2010 may be less rich than in previous or subsequent years, but this is certainly not because officers drafting IAs lowered their standards before the election. Many of the stronger months in this period actually fall in the period December 2009 to April 2010; in fact, March 2010 is the only month that is below the annual average. Instead, the low 2010 performance is clearly due to IAs prepared after the Conservative-Liberal government began to serve. We only have four IAs between June and September 2010 (and they were good), but when the newly elected government began to adopt a larger number of policies in October 2010, their IAs were always below the annual average. In other words, it is the post-election IAs that dragged the year down. There is a clear effect of dawn regulations. Perhaps not in the sense that policy-makers tried to get some pet projects out of the door immediately after the election. However, it is plausible that officers were somewhat insecure as to whether the newly elected government would place as much emphasis on the regulatory reform agenda, including regulatory analysis, as the previous Labour government did. Plus, the new government led by David Cameron was the first coalition government in more than 60 years; unsurprisingly, this had an impact on how well the administration operated after the new government assumed power.

The findings can, to a large degree, be explained by the context in which the two elections were contested: when the 2010 general elections were held, Labour had been in power for more than 13 years, and there was little confidence that Prime Minister Gordon Brown would succeed in winning the fourth subsequent victory since 1997 for his party. In other words, there was a realistic chance for change of government in 2010, and we are not surprised that, under these conditions, the logic of midnight and dawn regulations applies, although much less so than expected for midnight regulations. The 2005 elections, by contrast, were much less contested; the overall expectation that Prime Minister Tony Blair would land another victory for Labour may have ‘deactivated’ mechanisms usually resulting in midnight and dawn regulations. It is impossible to draw strong conclusions from these data, but it seems that elections have an influence, although not in the mechanical way that the simplistic midnight-dawn regulations argument suggests.

HYPOTHESIS 5: SPECIALIZATION IN POLICY DOMAINS

Let us now consider the argument that departments do not perform analysis evenly. The richness is deeper in sections that mirror their core mission and their stakeholders’ preferences.

To probe this, we consider four types of departments: economy-oriented, environmental policy-oriented, social policy-oriented and a residual category with the others. The underlying assumption is that departments operating, say, in the field of environment have their constituency there – such as the Department for Environment, Food and Rural Affairs – and invest more time and effort on the environmental impacts of a proposed policy than, for instance, the Department for Work and Pensions or the Cabinet Office. This generates a four-level categorical independent variable, summarized in Table 9 below.

--- TABLE 9 ABOUT HERE ---

It goes without saying that we cannot use the overall index to test this hypothesis. After all, we do not want to know whether departments specializing in environmental policies produce richer IAs than departments with a health and welfare portfolio. Statistically, what we want to know is whether IA requirements related to social dimensions, such as gender equality or access to health and education, are implemented more rigorously in a department specializing in social policy. Likewise, do departments like Business, Innovation and Skills analyze the impacts of their policies on trade, jobs or growth better than other departments?

To answer this question, we use sub-indexes 2, 3 and 4, aggregating data on the richness of IA related to economic, social and environmental impacts, respectively. We compare the mean values of these sub-indexes across different types of departments (economic, social, environmental and other specialization) using three separate one-way ANOVAs, one for each index. The one-way ANOVA on the sub-index on economic analyses is not significant. On the other hand, the two one-way ANOVAs on the social and environmental sub-indexes are significant, overall ($p<.01$ for both). Table 10 shows the mean values for each sub-index across each category of departments.

--- TABLE 10 ABOUT HERE ---

What does the data tell us? Specialist departments usually produce better analyses in their home category than other departments: the richest analyses of environmental impacts are carried out by environmental regulators; departments working on welfare, health, pensions submit the richest analyses of social impacts, and economic departments deliver above-average analyses of impacts on the economy.

In order to qualify the overall significance of the two ANOVAs on the social and environmental sub-indexes, we perform post-hoc multiple comparisons tests. With regards to the environmental sub-index, we observe a remarkable effect size, measured by Cohen's d, whereby the difference between the mean values of analytical richness of environmental analyses carried out by environmental departments is larger than one standard deviation with respect to the mean values of economic departments.[viii] Likewise, we observe a noticeable effect size value with regards to the social sub-index.[ix]

However, the story does not end here. This is because there are some effects that should not be there. According to our reasoning, departments do not specialize in categories outside their home category. Under this qualification, we are only able to confirm our intuition for environmental and social impacts. The findings on economic impacts are less straightforward. True, economic regulators produce above-average analyses of economic impacts. But in contrast to our expectation, those departments are not the only ones to return high-quality analyses of impacts on the economy. Environmental and social regulators are strong performers too when it comes to impacts on trade, growth and jobs.

To sum up then, departments learn more in their 'core business', but we also observed instances of specialization not predicted by this argument. We know little about the reasons why all departments do extraordinarily well when it comes to economic analyses. More qualitative research is needed. We offer two intuitions: first, in the age of austerity, environmental and social regulators are under exceptional pressure to justify further intervention into the economy, resulting in overall strong regulatory analysis. Second, governments have put a high premium on fostering economic growth – the official documentation instructs public managers to design policy having regard to growth and to the de-regulatory targets (Department for Business Innovation and Skills 2015). Consequently, analyses of economic impacts may have become a priority also for those departments regulating policy areas other than business and trade.

Although we observe departmental specialization, we know little about the underlying mechanisms. We offer three possible micro-causes: pressure of department-specific interest groups (this is how we framed the hypothesis); analytical capacity developed via intimate knowledge of regulated sectors; and finally, sector-specific analytical guidelines. Should expertise and knowledge explain departmental specialization, then we would ask whether there are significant differences between countries with different administrative cultures, specifically between countries (such as the UK) emphasizing the importance of generalists and those conceiving the public service as a realm of specialists (Bulmer 1988; Knill 2001). All this is material for further research.

HYPOTHESIS 6: WHERE YOU ARE IN THE EXECUTIVE MATTERS

Let us now reason that core departments, tasked with designing broad government policy agendas and controlling public expenditure (for instance the Cabinet Office or the Treasury), are more likely to support regulatory analysis, for example because they believe that IA really pre-empts inefficient regulation. Both Labour governments and the Coalition government have made efforts to increase control capacity at the centre (Dommett and Flinders 2015; Evans 2009). Classic regulatory departments like the Health and Safety Executive or the Department for Work and Pensions should be less enthusiastic about having to set aside precious resources for regulatory analysis and presenting their numbers in cabinet-level committees. An alternative proposition, however, might well be that regulatory departments deliver IAs that are broader and deeper because they find this is the best way to justify their task expansion.

To test these propositions, we distinguish three types of departments, thereby creating a new categorical grouping variable: first, core departments tasked to manage classic state functions such as foreign affairs, home policy, justice and the budget; second, peripheral departments associated with spending and regulation in areas like the environment, health, social welfare and education; third, finally, departments responsible for advancing the regulatory reform agenda in the UK. The assignment of departments follows suggestions made in the literature on the core executive (Dunleavy and Rhodes 1990; Elgie 2011). Table 11 below informs about the types of departments in our sample:

--- TABLE 11 ABOUT HERE ---

We then performed a one-way ANOVA to test the differences in the mean values of our index of analytical richness across the above categories. The mean values across typology of departments are presented in Table 12 below:

--- TABLE 12 ABOUT HERE ---

The findings, supported by further post-hoc multiple comparisons, suggest that peripheral departments outperform core departments with an average to low effect size[x] and also, quite surprisingly, BERR-BIS-DTI, though the latter result is not statistically significant. BERR-BIS-DTI is a department which has changed its name (and some tasks) multiple times, but has always been at the forefront when it comes to promoting the Better Regulation agenda. Our data reject Hypothesis 6a which suggested that core departments and, even more so, BERR-BIS-DTI would spearhead regulatory analysis. Instead, we find support for Hypothesis 6b: it is the peripheral department that produces higher-quality IAs. A solid IA - we reason - is a necessary condition to gather consensus within cabinet-level committees with de-regulation preferences. Given the de-regulatory zeal of Gordon Brown and George Osborne, a strong IA is a good way to defend regulatory proposals generated by the 'periphery'.

CONCLUSION

Our evidence suggests that experience and analytical capacity in different departments play a role in the learning process. Guidelines that explain how to carry out IA in practice support the process, but moderately. A specialized, IA-focused regulatory oversight body matters, and this may explain why the government has increased the responsibilities of the RPC in the years following the period we examined in our project. Elections have different effects on regulatory analysis, depending on the context in which they are contested – but we only had two elections in our period. Departments specialize and reflect their core policy sectors, but some have greater analytical capacity overall. Peripheral departments seem to invest more in IA, arguably because they know that robust analysis is key to cabinet discussions in an era of austerity and concerns about 'red tape' and regulatory costs.

And yet, does it matter? It does: whether the objective is the political control of the bureaucracy or more evidence-based policy (see introduction), detailed information is crucial. If IAs are not informative, do not report on major cost and benefit categories or are silent on consultation, a necessary condition for evidence-based policy is missing. At the same time, stakeholders in fire-alarm scenarios will be largely disempowered and principals left in the dark as to whether they are actually in control of their agents (McCubbins, Noll and Weingast 1987). Thus, it is topical to understand and explain what generates richness over time. The role and scope of the RPC and the official guidelines can be fine-tuned to draw more information from the IA – the recent evolution of the RPC is indicative of the attempt to leverage IAs figures to meet de-regulatory objectives. Targets like 'one-in-one-out' and better regulation framework manuals are evidence that governments manipulate IA to pursue their regulatory reform priorities.

At the same time, the difference between the de-regulatory agenda of the core executive and the regulatory missions of some departments may create conflicts between reducing regulatory costs and protecting lives and the environment. Here the apparently technical exercise of performing one type of analysis or another becomes the terrain where regulatory policy paradigms may clash. This may be part of the broader story of whether parties in government have more influence on IA than the preferences of departments (in the UK) or agencies (in the US) – Ellig and colleagues (2013) looked at the US and identified ‘conservative agencies’ that produce better analysis in the Obama administration, and ‘liberal agencies’ that perform better during the Bush administration.

Methodologically, we strongly defend our choice of having appraised six potential factors with nuanced, tailored statistical analyses rather than opting for a classic multivariate model. Our ‘gentle approach’ to the data allows us to contribute to the key themes in the field about political control of the bureaucracy, learning and evidence-based policy with more detail and nuances. This – we believe – is also the type of research finding that is more useful to policy-makers.

Our results come with caveats: we need more research on a longer time-span covering the years after 2011. Also, regulatory analysis is a component of legislative and parliamentary processes, at least in the UK. The next generation of projects should regress our data on different variables, such as duration, conflict in the lawmaking process, media attention and possibly data generated by post-implementation reviews of regulations that were originally appraised via IA. Our index has properties that make it more suitable for research across time and space than other scorecards measures of IA content proposed in the past. Researchers could use it within sophisticated models of the policy agenda (Baumgartner and Jones 1993). Future research should also distinguish major and minor policy proposals; this is because departments in the UK are encouraged to take a lighter approach towards IA if the impacts are likely to be minor. Finally, qualitative and ethnographic researchers could extract from our data the richest and poorest IAs and document the different usages by bureaucrats and politicians in policy formulation. There is a whole story to tell about meanings, interpretations and usages of regulatory analysis, thus reconnecting studies like ours to the broader field of knowledge utilization.

References

Adelle C, Weiland S, Dick J, González Olivo D, Marquardt J, Rots G, Wübbeke J, Zasada I (2016) Regulatory Impact Assessment: a Survey of Selected Developing and Emerging Economies. *Public Money & Management 36*, 89-96.

Baumgartner FR, Jones BD (1993) *Agendas and Instability in American Politics*. Chicago: University of Chicago Press.

Beermann JM (2009) Combating Midnight Regulations. *Northwestern Law Review 103*, 352-369.

Belcore J, Ellig J (2008) Homeland Security and Regulatory Analysis: Are We Safe Yet? *Rutgers Law Journal 40*, 1-96.

Boswell C (2008) The Political Functions of Expert Knowledge: Knowledge and Legitimation in European Union Immigration Policy. *Journal of European Public Policy 15*, 471-488.

Bulmer M (1988) Science Expertise and Executive-Bureaucratic Politics in Britain. *Governance 1*, 26-49.

Cabinet Office (2003) Regulatory Impact Assessment Guidance. The Stationery Office: London.

Cairney P (2015) *The Politics of Evidence-Based Policymaking*. Basingstoke: Palgrave.

Canelas L, Almansa P, Merchan M, Cifuentes P (2005) Quality of Environmental Impact Statements in Portugal and Spain. *Environmental Impact Assessment Review 25*, 217-225.

Carley M (1980) *Rational Techniques in Policy Analysis*. London: Heinemann Educational.

Carpenter DP (2010) *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton: Princeton University Press.

Carpenter DP, Gubb J (2014) A Political Clearinghouse? Organizational and Presidential Effects in OIRA Review. In *Midwest Political Science Association (MPSA)*. Chicago.

Carroll P (2010) Does Regulatory Impact Assessment Lead to Better Policy? *Policy and Society 29*, 113-122.

Cecot C, Hahn R, Renda A, Schrefler L (2008) An Evaluation of the Quality of Impact Assessment in the European Union with Lessons for the US and the EU. *Regulation & Governance 2*, 405-424.

Damonte A, Dunlop CA, Radaelli CM (2014) Controlling Bureaucracies with Fire Alarms: Policy Instruments and Cross-Country Patterns. *Journal of European Public Policy 21*, 1330-1349.

De Francesco F (2012) Diffusion of Regulatory Impact Analysis in OECD and EU Member States. *Comparative Political Studies 45*, 1277-1305.

Department for Business Enterprise and Regulatory Reform (2007) Impact Assessment Guidance.

Department for Business Innovation and Skills (2015) Better Regulation Framework Manual: Practical Guidance for UK Government Officials. March 2015. London.

Desmarais BA, Hird JA (2014) Public Policy's Bibliography: the Use of Research in US Regulatory Impact Analyses. *Regulation & Governance 8*, 497-510.

Dommett K, Flinders M (2015) The Centre Strikes back: Meta-Governance, Delegation, and the Core Executive in the United Kingdom, 2010-14. *Public Administration 93*, 1-16.

Downs A (1966) *Inside Bureaucracy*. Boston: Little & Brown.

Dunleavy P, Rhodes RAW (1990) Core Executive Studies in Britain. *Public Administration 68*, 3-28.

Dunlop C, Radaelli C, eds (2016a) *Handbook of Regulatory Impact Assessment*. Cheltenham: Edward Elgar.

Dunlop CA, Radaelli CM (2016b) Policy Learning and Organizational Capacity. In *The Palgrave Handbook of Public Administration and Public Management in Europe*: Palgrave.

Elgie R (2011) Core Executive Studies Two Decades on. *Public Administration 89*, 64-77.

Ellig J, McLaughlin PA (2012) The Quality and Use of Regulatory Analysis in 2008. *Risk Analysis 32*, 855-880.

Ellig J, McLaughlin PA, Morrall III JF (2013) Continuity, Change, and Priorities: the Quality and Use of Regulatory Analysis across US Administrations. *Regulation & Governance 7*, 153-173.

Ellig J, McLaughlin PA, Morrall JF (2013) Continuity, Change, and Priorities: the Quality and Use of Regulatory Analysis across US Administrations. *Regulation & Governance 7*, 153-173.

Evans M (2009) Gordon Brown and Public Management Reform – a Project in Search of a 'Big Idea'? *Policy Studies 30*, 33-51.

Fraas A, Lutter RW (2013) The Challenges of Improving the Economic Analysis of Pending Regulations: the Experience of OMB Circular A-4. *Annual Review of Resource Economics 3*, 71-85.

Fritsch O, Radaelli C, Schrefler L, Renda A (2013) Comparing the Content of Regulatory Impact Assessments in the UK and the EU. *Public Money & Management 33*, 445-452.

Froud J, Boden R, Ogus A, Stubbs P (1998) *Controlling the Regulator*. Basingstoke: Macmillan.

Gibbons M, Parker D (2012) Impact Assessments and Better Regulation: the Role of the UK's Regulatory Policy Committee. *Public Money & Management 32*, 257-264.

Gibbons M, Parker D (2013) New Development: Recent Changes to the UK's Regulatory Process. *Public Money & Management*

Gray I, Edward-Jones G (2003) A Review of Environmental Statements in the British Forest Sector. *Impact Assessment and Project Appraisal 21*, 303-312.

Hahn RW, Tetlock PC (2008) Has Economic Analysis Improved Regulatory Decisions? *Journal of Economic Perspectives 22*, 67-84.

Hertin J, Turnpenny J, Jordan A, Nilsson M, Russel D, Nykvist B (2009) Rationalising the Policy Mess? Ex Ante Policy Assessment and the Utilisation of Knowledge in the Policy Process. *Environment and Planning A 41*, 1185-1200.

HM Government (2011) One-in, One-out: Statement of New Regulation. April 2011.

Kabir SMZ, Momtaz S (2014) Sectorial Variation in the Quality of Environmental Impact Statements and Factors Influencing the Quality. *Journal of Environmental Planning and Management 57*, 1595-1611.

Kelman S (2005) *Unleashing Change: a Study of Organizational Renewal in Government*. Washington: Brookings Institutions Press.

Knill C (2001) *The Europeanisation of National Administrations: Patterns of Institutional Change and Persistence*. Cambridge: Cambridge University Press.

Landim SNT, Sánchez LE (2012) The Contents and Scope of Environmental Impact Statements: How Do They Evolve over Time? *Impact Assessment and Project Appraisal 30*, 217-228.

Lodge M, Wegrich K (2009) High-quality regulation: its popularity, its tools and its future. *Public Money & Management 29*, 145-152.

Lodge M, Wegrich K (2015) Crowdsourcing and Regulatory Reviews: a New Way of Challenging Red Tape in British Government? *Regulation & Governance 9*, 30-46.

McCubbins MD, Noll RG, Weingast BR (1987) Administrative Procedures as Instruments of Political Control. *Journal of Law, Economics, & Organization 3*, 243-277.

McGarity TO (1991) *Reinventing Rationality: the Role of Regulatory Analysis in the Federal Bureaucracy*. Cambridge: Cambridge University Press.

Nutley SM, Walter I, Daviers HTO (2007) *Using Evidence: How Research Can Inform Public Services*. Bristol: Policy Press.

OECD. 2009. Indicators of Regulatory Management Systems. Paris.

Ogus A (2001) *Regulation, Economics, and the Law*. Cheltenham: Edward Elgar.

Peci A, Sobral F (2011) Regulatory Impact Assessment: How Political and Organizational Forces Influence Its Diffusion in a Developing Country. *Regulation & Governance 5*, 204-220.

Pinho P, Rodrigo M, Monterrosso A (2007) The Quality of Portuguese Environmental Impact Studies: the Case of Small Hydropower Projects. *Environmental Impact Assessment Review 27*, 189-205.

Prosser T (2010) *The Regulatory Enterprise: Government, Regulation, and Legitimacy*. Oxford: Oxford University Press.

Radaelli CM (2005) Diffusion without Convergence: How Political Context Shapes the Adoption of Regulatory Impact Assessment. *Journal of European Public Policy 12*, 924-943.

Regulatory Policy Committee (2010) Reviewing Regulation: an Independent Report on the Analysis Supporting Regulatory Proposals, December 2009-May 2010. London. August 2010.

Renda A (2011) *Law and Economics in the RIA World: Improving the Use of Economic Analysis in Public Policy and Legislation*. Cambridge: Intersentia.

Russel D, Jordan A (2009) Joining up or Pulling Apart? The Use of Appraisal to Coordinate Policy Making for Sustainable Development. *Environment and Planning A 41*, 1201-1216.

Seltman HJ (2015) *Experimental Design and Analysis*. Pittsburgh: Carnegie Mellon University.

Shapiro S, Morrall JF (2012) The Triumph of Regulatory Politics: Benefit-Cost Analysis and Political Salience. *Regulation & Governance 6*, 189-206.

Shapiro S, Morrall JF (2016) Does Haste Make Waste? How Long Does It Take to Do a Good Regulatory Impact Analysis? *Administration & Society 48*, 367-389.

Smismans S, Minto R (2016) Are Integrated Impact Assessments the Way forward for Mainstreaming in the European Union? *Regulation & Governance*

Staronová K (2010) Regulatory Impact Assessment: Formal Institutionalization and Practice. *Journal of Public Policy 30*, 117-136.

Tzoumis K (2007) Comparing the Quality of Draft Environmental Impact Statements by Agencies in the United States since 1998 to 2004. *Environmental Impact Assessment Review 27*, 26-40.

Tzoumis K, Finegold L (2000) Looking at the Quality of Draft Environmental Impact Statements over Time in the United States: Have Ratings Improved? *Environmental Impact Assessment Review 20*, 557-578.

Vyas S, Kumaranayake L (2006) Constructing Socio-Economic Status Indices: How to Use Principal Components Analysis. *Health Policy and Planning 21*, 459-468.

Wiener JB (2007) Better Regulation in Europe. *Current Legal Problems 59*, 447-518.

Wilson JQ (1991) *Bureaucracy: What Government Agencies Do and Why They Do It*: Basic Books.

Figure 1. Scatterplot of the values of the main index. The index (y-axis) ranges from 0 to 1. IAs are grouped according to the month and year of publication (x-axis). The blue line represents the linear fit line (R2 linear =.035). The red line represents the locally weighted scatterplot smoothing line (LOESS: Epanechnikov Kernel, 50 per cent of points fitted).

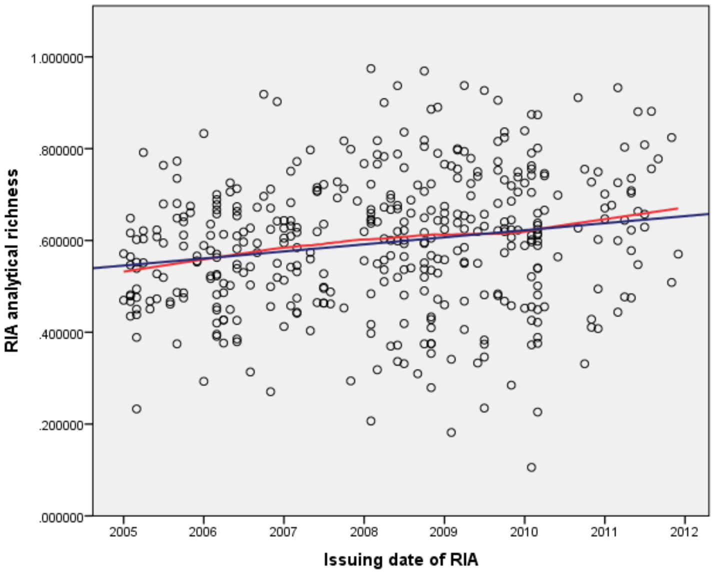


Table 1. Analytical richness of IA per department and number of IAs, based on the main index. The index ranges from 0 to 1.

DEPARTMENT	DEFRA	DH	DT	DTI	FSA	HMRC	HMT	HO
Number of IAs[xi]	56	28	45	30	25	35	21	21
First 5 IAs	0.528	0.522	0.631	0.491	0.447	0.572	0.538	0.641
First 10 IAs	0.577	0.544	0.614	0.522	0.524	0.532	0.498	0.645
First 15 IAs	0.581	0.569	0.621	0.511	0.569	0.530	0.515	0.624
First 20 IAs	0.565	0.564	0.624	0.502	0.577	0.529	0.553	0.622
First 25 IAs	0.576	0.560	0.618	0.514	0.569	0.552		
First 30 IAs	0.572		0.616	0.533		0.573		
First 40 IAs	0.589		0.624					
First 50 IAs	0.617							

Table 2. T-test, mean values of the sub-index on cost-benefit analysis before and after the publication of the 2007 IA guideline. The index ranges from 0 to 1.

DATE	MEAN	STD DEVIATION
Pre 2007 guidelines	0.464	0.149
Post 2007 guidelines	0.596	0.184
p<.01; Cohen's d=.788		

Table 3. T-test, mean values of the IA index before and after the publication of the 2007 IA guideline. The index ranges from 0 to 1.

DATE	MEAN	STD DEVIATION
Pre 2007 guidelines	0.559	0.124
Post 2007 guidelines	0.608	0.160
$p<.01$; Cohen's d=.342		

Table 4. T-test, mean values of the IA index before and after the establishment of the RPC. The index ranges from 0 to 1.

DATE	MEAN	STD DEVIATION
Pre RPC	0.581	0.145
Post RPC	0.612	0.159
p=.027; Cohen's d=.204		

Table 5: Contrast coefficients.

YEAR	2005	2006	2007	2008	2009	2010	2011
Contrast	-1	1	0	0	1	-2	1

Table 6: Mean values of the IA index across years. The index ranges from 0 to 1.

YEAR	N	MEAN	STD DEVIATION
2005	82	0.551	0.127
2006	93	0.558	0.127
2007	73	0.598	0.131
2008	96	0.590	0.164
2009	78	0.618	0.154
2010	56	0.586	0.168
2011	37	0.650	0.147

Table 7. IA index before and after the general election of 5 May 2005. The index ranges from 0 to 1.

MONTH	NOV 2004	DEC 2004	JAN 2005	FEB 2005	MAR 2005	APR 2005	MAY 2005	TOTAL
IA INDEX (MIDNIGHT)	N/A	N/A	0.520	0.530	0.464	0.642	0.459	0.554
MONTH	JUN 2005	JUL 2005	AUG 2005	SEP 2005	OCT 2005	NOV 2005	DEC 2005	
IA INDEX (DAWN)	0.558	0.639	0.464	0.616	0.575	0.668	0.558	

Table 8. Quality of RIA before and after the general election of 6 May 2010. The index ranges from 0 to 1.

MONTH	**NOV 2009**	**DEC 2009**	**JAN 2010**	**FEB 2010**	**MAR 2010**	**APR 2010**	**MAY 2010**	**TOTAL**
IA INDEX (MIDNIGHT)	0.576	0.704	0.651	0.594	0.556	0.652	NONE	0.589
MONTH	**JUN 2010**	**JUL 2010**	**AUG 2010**	**SEP 2010**	**OCT 2010**	**NOV 2010**	**DEC 2010**	
IA INDEX (DAWN)	0.595	NONE	NONE	0.769	0.543	0.522	0.563	

Table 9: Departments across policy sectors.

POLICY SECTOR	DEPARTMENTS
Economic[xii]	BERR, BIS, DFES, DIUS, DTI, HMT
Environmental	DECC, DEFRA, FC
Social	DCFS, DH, DWP, FSA, HSE
Other	CO, DCA, DCLG, DCMS, DT, FCO, HMRC, HO, MOJ, ODPM

Table 10: Indexes of sectoral analyses across departments. The indexes range from 0 to 1.

DEPARTMENTAL SPECIALIZATION	ECONOMIC IMPACTS	SOCIAL IMPACTS	ENVIRONMENTAL IMPACTS
Economic	0.652	0.502	0.402
Environmental	0.648	0.492	0.519
Social	0.633	0.585	0.423
Other	0.645	0.547	0.442

Table 11: Core and non-core departments in the UK.

TYPE OF DEPARTMENT	DEPARTMENTS
Regulatory reform	BERR, BIS, DTI
Core	CO, DCA, DIUS, FCO, HMRC, HMT, HO, MOJ, ODPM
Peripheral	DCLG, DCMS, DCSF, DECC, DEFRA, DFES, DH, DT, DWP, FC, FSA, HSE

Table 12: One-way ANOVA of IA index across core, peripheral and reform-oriented departments. The index ranges from 0 to 1.

DEPARTMENTS	MEAN	STD DEVIATION
Regulatory reform	0.574	0.14
Core	0.560	0.148
Peripheral	0.602	0.147
ANOVA overall significance p=.019		

SUPPLEMENTARY MATERIALS

Table A1. List of departments

ACRONYM	DEPARTMENT	OPERATING
BERR	Department for Business, Enterprise and Regulatory Reform	2007 - 2009
BIS	Department for Business, Innovation and Skills	2009 -
CO	Cabinet Office	1916 -
DCA	Department for Constitutional Affairs	2003 - 2007
DCLG	Department for Communities and Local Government	2006 -
DCMS	Department for Culture, Media and Sport	1997 -
DCSF	Department for Children, Schools and Families	2007 - 2010
DECC	Department for Energy and Climate Change	2008 -
DEFRA	Department for Environment, Food and Rural Affairs	2001 -
DFES	Department for Education and Skills	2001 - 2007
DH	Department of Health	1988 -
DIUS	Department for Innovation, Universities and Skills	2007 - 2009
DT	Department for Transport	2002 -
DTI	Department of Trade and Industry	1970 - 2007
DWP	Department for Work and Pensions	2001 -
FC	Forestry Commission	1919 -
FCO	Foreign and Commonwealth Office	1968 -
FSA	Food Standards Agency	2000 -
HMRC	Her Majesty's Revenues and Customs	2005 -
HMT	Her Majesty's Treasury	1066 -
HO	Home Office	1782 -
HSE	Health and Safety Executive	1974 -
MOJ	Ministry of Justice	2007 -
ODPM	Office of the Deputy Prime Minister	2001 - 2006

Table A2. Number of IAs in our sample, per department and year

DEPARTMENTS	2005	2006	2007	2008	2009	2010	2011	TOTAL
BERR	--	--	2	13	4	--	--	19
BIS	--	--	--	--	4	6	5	15
CO	0	0	0	1	0	0	1	2
DCA	1	2	--	--	--	--	--	3
DCLG	0	6	9	19	8	11	10	63
DCMS	3	4	5	1	3	4	0	20
DCSF	--	--	0	2	0	0	--	2
DECC	--	--	--	0	3	4	4	11
DEFRA	13	12	14	14	9	3	2	67
DFES	4	1	1	--	--	--	--	6
DH	8	4	4	10	5	1	0	32
DIUS	--	--	1	1	1	--	--	3
DT	14	6	10	7	8	10	9	64
DTI	9	28	4	--	--	--	--	41
DWP	3	1	1	2	1	1	0	9
FC	0	1	0	0	0	0	0	1
FCO	1	0	1	0	1	0	0	3
FSA	5	4	7	3	3	3	0	25
HMRC	6	7	7	8	8	3	1	40
HMT	1	6	1	4	5	1	4	22
HO	5	4	3	10	6	2	1	31
HSE	3	3	0	1	2	0	0	9
MOJ	0	0	3	2	7	7	0	20
ODPM	6	4	--	--	--	--	--	10
Total	82	93	73	98	78	56	37	517

Table A3. Scorecard

NO	NAME
THE SCORED IA	
1	First name of scorer
2	Date of scoring
3	Time required
4	Name of policy initiative
5	Origin of policy initiative
6	Type of policy initiative
7	Department or agency preparing IA
8	Joint submission of IA
9	Year of publication
10	Number of pages
11	Summary page
PROBLEM IDENTIFICATION	
12	Identifies market failure
13	Identifies regulatory failure
14	States objectives
15	States specific objectives
16	States operational objectives
POLICY OPTIONS	
17	Considers the zero option
18	Considers at least one alternative to the zero option
19	Considers at least two alternatives to the zero option
20	Considers improvements in implementation and enforcement
21	Considers self-regulation
22	Considers regulation through information and guidelines
23	Considers regulation through market-based instruments
24	Considers regulation through direct public sector financial intervention
25	Considers co-regulation
26	Considers prescriptive regulatory actions
CONSULTATION	
27	Reports on consultation

28	Presents positions expressed by consulted parties
29	Cooperation between departments
ESTIMATION OF COSTS AND BENEFITS OF THE SUGGESTED POLICY OPTION	
30	Presents qualitative or quantitative statements on costs
31	Quantifies at least some costs
32	Monetizes at least some costs
33	Monetizes all or nearly all costs
34	Provides range for total costs
35	Presents qualitative statements on benefits
36	Quantifies at least some benefits
37	Monetizes at least some benefits
38	Monetizes all or nearly all benefits
39	Provides range for total benefits
40	Calculates net benefits
41	Provides a range for net benefits
42	Calculates cost effectiveness
ESTIMATION OF COSTS AND BENEFITS OF ALTERNATIVE POLICY OPTIONS	
43	Presents qualitative statements on costs of at least one alternative option
44	Quantifies at least some costs of at least one alternative option
45	Monetizes at least some costs of at least one alternative option
46	Monetizes all or nearly all costs of all options
47	Provides range for total costs of at least one alternative option
48	Presents qualitative statements on benefits of at least one alternative option
49	Quantifies at least some benefits of at least one alternative option
50	Monetizes at least some benefits of at least one alternative option
51	Monetizes all or nearly all benefits of all options
52	Provides range for total benefits of at least one alternative option
53	Calculates net benefits of at least one alternative option
54	Provides a range for net benefits of at least one alternative option
55	Calculates cost effectiveness of at least one alternative option
ANALYSES	
56	Carries out risk assessment
57	Carries out risk-risk analysis

58	Considers precautionary principle
59	Carries out sensitivity analysis
60	Identifies discount rate
61	Value of discount rate
62	Provides number of lives or of life years or quality-adjusted life-years (QUALYs) saved
63	Monetizes number of lives saved
AFFECTED PARTIES	
64	Discusses whether regulation imposes costs on citizens
65	Discusses whether regulation imposes costs on specific categories of citizens
66	Discusses whether regulation imposes costs on consumers
67	Discusses whether regulation imposes costs on the economic sector
68	Discusses whether regulation imposes costs on a few large firms
69	Discusses whether regulation imposes costs on the non-profit sector
70	Discusses whether citizens benefit from regulation
71	Discusses whether specific categories of citizens benefit from regulation
72	Discusses whether consumers benefit from regulation
73	Discusses whether the economic sector benefits from regulation
74	Discusses whether a few large firms benefit from regulation
75	Discusses whether the non-profit sector benefits from regulation
ECONOMIC IMPACTS	
76	Assesses impact on competitiveness
77	Assesses impact on competition
78	Assesses impact on small and medium enterprises
79	Assesses impact on investment or innovation
80	Assesses impact on the common market
81	Assesses impact on GDP or other indicators of economic growth
82	Assesses impact on trade
83	Assesses impact on inflation
84	Assesses impact on administrative burdens
85	Quantifies administrative burdens for businesses
86	Quantifies administrative burdens for citizens
87	Quantifies administrative burdens for public administration
SOCIAL IMPACTS	

88	Assesses impact on health and safety
89	Assesses impact on employment
90	Assesses impact on standards and rights related to job quality
91	Assesses impact on the social inclusion and protection of particular groups
92	Assesses impact on equal opportunities, non-discrimination and gender equality
93	Assesses impact on the access to and effects on social protection, health and education
94	Assesses impact on fundamental rights
ENVIRONMENTAL IMPACTS	
95	Assesses impact on renewable or non-renewable resources
96	Assesses impact on biodiversity
97	Assesses impact on air quality
98	Assesses impact on transport and the use of energy
99	Assesses impact on water quality
100	Assesses impact on soil quality and resources
101	Assesses impact on climate
MONITORING AND EVALUATION	
102	Contains a section on monitoring and evaluation
103	Mentions a review clause for the proposal
104	Contains indicators for evaluation
OTHER	
105	Overall judgment
106	Additional comments
107	File name

PRINCIPAL COMPONENT ANALYSIS (INCLUDING TABLES A4, A5 AND A6 AND FIGURE A1)

Our coding relied on a classic scorecard approach, widely used by policy makers and academics. We then used Principal Component Analysis (PCA) to aggregate our data, this step reflects an inductive approach towards our data. This section motivates, details and provides supporting statistics of the various steps undertaken during the PCA.

The degree to which IAs comply with guideline requirements is not a latent trait which can be described by an underlying hidden model or a path. We therefore chose a dimension reduction technique that enables us to summarise our data without reference to a specific model. Instead, the technique relies on a simple idea: to explain the total (maximised) variability of our sample.[1] Furthermore, thanks to PCA, or similar dimension reduction techniques such as Multiple Correspondence Analysis, we avoid imposing a fixed number of factors to be extracted. This is an advantage as compared to Exploratory Factor Analysis. Finally, a sufficient number of bivariate correlations between scorecard items are statistically significant (two-tailed significance) and sizeable (i.e. above the .3 threshold), another reason to run a PCA based on the correlation matrix.

[1] "[F]actor analysis attempts to achieve a reduction from p [manifest variables] to m dimensions by invoking a *model* relating $x_1, x_2, \dots, x_p$ to m hypothetical or latent variables […] PCA differs from factor analysis in having *no explicit model*." (Joliffe 2002, p. 151, emphasis in original).

We considered the following variables in our PCA: 12 to 42; 56 to 60; 62-104. Total: 79 variables.[2]

The literature suggests using either the simple Pearson's correlation matrix or the tetrachoric correlation matrix. We rely on the simple correlation matrix. Two reasons: first, we found the tetrachoric correlation matrix too difficult to calculate – in terms of computer power, i.e. the solution did not converge. After all, we are speaking of a 79x79 correlation matrix across 517 observations. Second, the key condition for the use of tetrachoric correlations - latent bivariate normality - is not met in our case. As we will explain later below, further diagnostics confirm the validity of our dimension reduction based on the Pearson's correlation matrix.

In a next step, we chose the most appropriate rotation technique. Rotations may be either orthogonal or oblique. Key factors to keep in mind here are the ex-post detection of a simple structure and a theoretically supported expectation as to whether the extracted components will be correlated or not. We performed two rotations, one oblique (Direct Oblimin, $\delta = 0$) and one orthogonal (Varimax) with a view to compare the outcomes. Initially, we were slightly more lenient towards an orthogonal Varimax rotation because it enabled us to maximize the variance. The aim of the PCA is, after all, to describe parsimoniously the key sources of variability among our observed variables. The Component Correlation Matrix, a result of the Direct Oblimin rotation of the robust PCA iteration (see below), suggested that the components were loosely correlated to each other (no correlation above the .3

2 See Table A3 above for our scorecard.

threshold), indicating that the oblique solution closely approached orthogonality. Furthermore, the results of the Varimax rotation on the robust PCA iteration, performed in parallel with the Oblimin rotation, suggested the presence of a simple structure underlying our components (see Table A5 below). We therefore decided to use the results of the Varimax rotation in subsequent analyses.

In order to ensure the robustness and appropriateness of the PCA, we applied several consistency criteria. In a nutshell, we ran three iterations of the PCA: first, on the full dataset. Second, on the full dataset but excluding variables with a communality below .5. The communality is the share - expressed on 0-1 scale - of the variance of each variable that may be explained by the extraction of those principal components. If the coefficient was below .5, we removed the variable and reran the PCA without it. Third, further excluding variables with a complex structure. A complex structure can be observed when a variable shows a significant loading - i.e. >0.4 - on more than one component. When we identified such a variable, we removed it from the set of manifest variables and reran the PCA without it.

When we ran those iterations, we carried out a number of tests in order to check for sample adequacy. In iteration 1, the Kaiser-Meyer-Olkin measure resulted in a value of 0.722, this is larger than the 0.6 usually suggested in the literature. The Bartlett's test of sphericity was significant at the 0.01 level. The anti-image correlation also confirmed the adequacy of our sample; to this end, we observed the coefficients on the diagonal axis of the matrix - they were all >0.5. We then analysed the ‘Communalities’ chart. In line with our criteria for exclusion (see above), we removed the variables 24, 83, 86 and 93 and reiterated the analysis without them. - In the second iteration, the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity still produced adequate results: the Kaiser-Meyer-Olkin measure was 0.723, and Bartlett's test was significant at the 0.01 level. The anti-image correlation also confirmed the adequacy of our sample. The communalities were all >0.5. We therefore proceeded to analyse the loadings of variables on the principal components to verify the presence of a simple structure. To this end, we looked at the Rotated Component Matrix (rotation converged in 28 iterations). Accordingly, the variables 30, 35, 38, 85 and 101 showed a complex structure, i.e. they have a loading >0.4 on more than one component. We therefore ran a third iteration of the PCA without those variables. This iteration demonstrated that variable 59 had a communality coefficient slightly below .5 (.498), hence we proceeded to a fourth iteration without it. Iteration four suggested that variables 29, 70 and 81 have low communalities, we therefore removed them and ran fifth PCA iteration. In iteration five all communalities were consistent. The Rotated Component Matrix (Table A5) highlighted that no component showed a complex structure (rotation converged in 37 iterations). Consequently, we relied on this PCA iteration in all subsequent analyses. The Kaiser-Meyer-Olkin measure (.704) and Bartlett's test of sphericity (significant at the .01 level) were adequate, as were the test on the anti-image matrix.

Finally, we performed another test to confirm the consistency of the PCA in its fifth iteration: we compared the outcomes of this iteration with those of two randomised validation samples extracted from our database. The results suggest that our fifth iteration was consistent with the two validation samples. Furthermore, analysing the correlation matrix for multicollinearity would be an option in a context of Confirmatory or Exploratory Factor Analysis. However, because we are dealing with a PCA based on orthogonal rotation, multicollinearity is not an issue.

Let us now have a look at the components. Table A4 below shows the total variance explained by the 24 principal components, retained according to the Kaiser criterion, i.e. eigenvalue > 1.

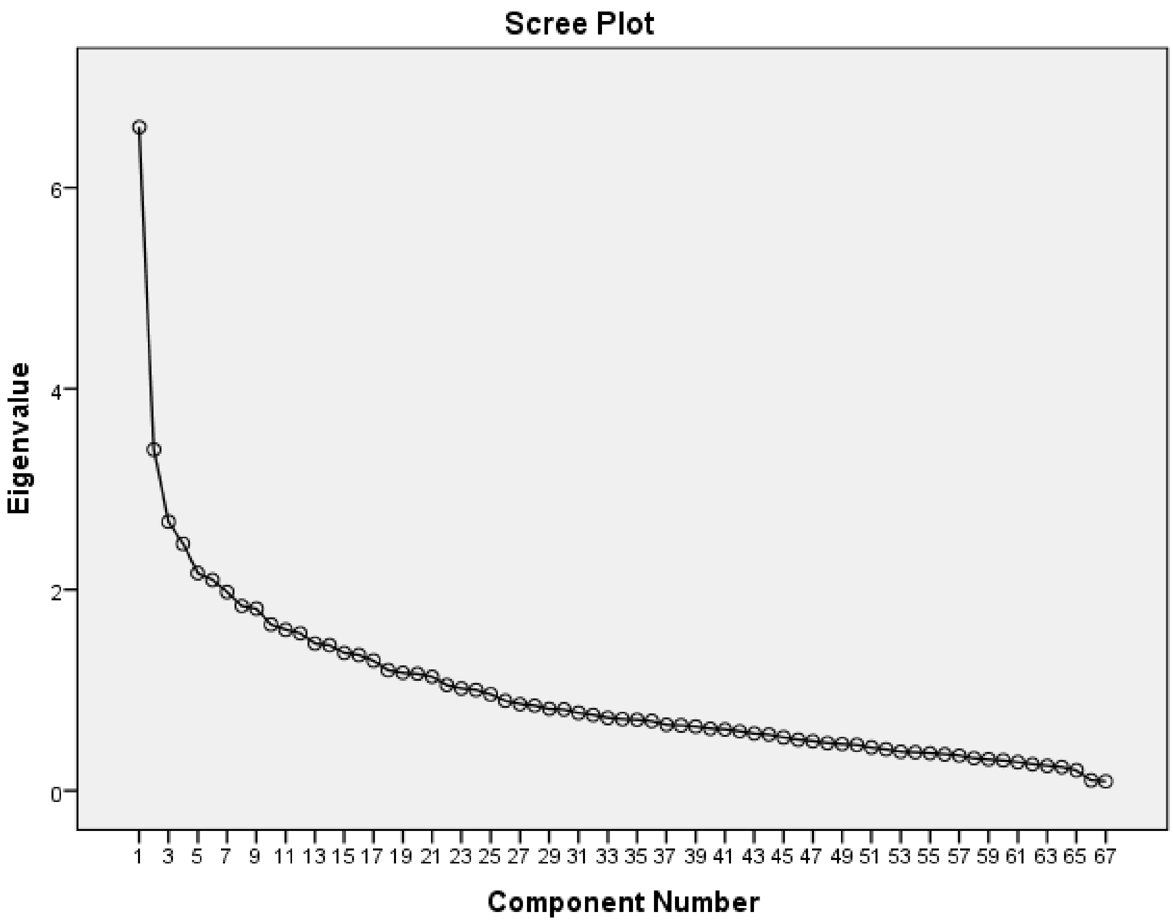
Table A4. Total Variance Explained by principal components

	Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	6.603	9.856	9.856
2	3.395	5.067	14.922
3	2.676	3.994	18.916
4	2.457	3.668	22.584
5	2.165	3.232	25.816
6	2.098	3.132	28.948
7	1.979	2.954	31.902
8	1.839	2.745	34.647
9	1.811	2.703	37.350
10	1.654	2.469	39.819
11	1.603	2.393	42.212
12	1.567	2.339	44.551
13	1.466	2.188	46.739
14	1.448	2.162	48.901
15	1.372	2.047	50.948
16	1.352	2.018	52.966
17	1.294	1.932	54.897

18	1.200	1.791	56.688
19	1.174	1.753	58.440
20	1.164	1.738	60.178
21	1.135	1.694	61.873
22	1.052	1.570	63.442
23	1.018	1.519	64.961
24	1.003	1.497	66.459

Those 24 components explain almost two thirds of the total variance in our sample. Please note that each principal component is uncorrelated - i.e. orthogonal - to the others. In order to reduce the number of those components, we used a Scree Plot graph, identifying nine components followed by a ‘bend’, explaining roughly 37.5 per cent of the total variance (Figure A below).

Figure A1



Aiming to increase the explanatory power of the PCA, we decided to abide by the criterion of “more than 50% of explained variance” (Joliffe 2002) and retained the first 15 components. Those 15 components constitute the basis of our index and are the most important sources of variation of our dataset. Table A5 below displays the rotated component matrix. It provides evidence that all components have a simple structure (i.e. single variables significantly load on one component only).

Table A5. Rotated Component Matrix

The influence of manifest variables on a principal component is measured through their loading coefficient. A simple structure emerges from PCA when “each component has a small number of large loadings and a large number of zero (or small) loadings” (Abdi and Williams 2010, 442). Therefore, simple structures can be observed when a subset of conceptually related manifest variables significantly load on a single component, thereby confirming that these manifest variables did indeed measure the same construct which is now aptly summarized by the component (see Table A5 in the appendix). To guarantee robustness and increase interpretability we performed a varimax orthogonal rotation of the principal components.

- This table is available upon request from the authors. -

Based on this, Table A6 below provides details about the interpretation we gave to the 15 retained components:

Table A6. Principal components

COMPONENTS	CUMULATIVE SHARE OF EXPLAINED VARIANCE	LOADING VARIABLES AND COEFFICIENTS
1 Benefit-cost measures	9.856% (9.856%)	34 Provides range for total costs (.643) 39 Provides range for total benefits (.816) 40 Calculates net benefits (.674) 41 Provides a range for net benefits (.811) 42 Calculates cost effectiveness (.510)
2 Cost quantification / monetization	5.067% (14.992%)	31 Quantifies at least some costs (.822) 32 Monetizes at least some costs (.897) 33 Monetizes all or nearly all costs (.657)
3 Consultation	3.994% (18.916%)	27 Reports on consultation (.750) 28 Presents positions expressed by consulted parties (.763) 29 Cooperation between departments (.634)
4 Impacts on specific economic sectors	3.668% (22.584%)	67 Discusses whether regulation imposes costs on the economic sector (.774) 73 Discusses whether the economic sector benefits from regulation (.773)

5 Social impacts	3.232% (25.816%)	88 Assesses impact on health and safety (.463) 91 Assesses impact on the social inclusion and protection of particular groups (.538) 92 Assesses impact on equal opportunities, non-discrimination and gender equality (.737) 94 Assesses impact on fundamental rights (.720)
6 Environmental impacts I	3.132% (28.948%)	95 Assesses impact on renewable or non-renewable resources (.668) 97 Assesses impact on air quality (.642) 98 Assesses impact on transport and the use of energy (.723)
7 Economic impacts on large firms	2.954% (31.902%)	68 Discusses whether regulation imposes costs on a few large firms (.922) 74 Discusses whether a few large firms benefit from regulation (.924)
8 Impacts on specific categories of citizens	2.745% (34.647%)	65 Discusses whether regulation imposes costs on specific categories of citizens (.814) 71 Discusses whether specific categories of citizens benefit from regulation (.838)
9 Environmental impacts II	2.703% (37.350%)	96 Assesses impact on biodiversity (.606) 99 Assesses impact on water quality (.769) 100 Assesses impact on soil quality and resources (.648)
10 Monitoring and implementation	2.469% (39.819%)	102 Contains a section on monitoring and evaluation (.701) 103 Mentions a review clause for the proposal (.642) 104 Contains indicators for evaluation (.576)
11 Administrative burdens	2.393% (42.212%)	84 Assesses impact on administrative burdens (.758) 87 Quantifies administrative burdens for public administration (.791)
12 Impacts on consumers	2.339% (44.551%)	66 Discusses whether regulation imposes costs on consumers (.805) 72 Discusses whether consumers benefit from regulation (.823)
13 Macro-economic impacts	2.188% (46.793%)	76 Assesses impact on competitiveness (.490) 80 Assesses impact on the common market (.713) 81 Assesses impact on GDP or other indicators of economic growth (.494) 82 Assesses impact on trade (.629)
14 Risk-related analyses	2.162% (48.901%)	56 Carries out a risk assessment (.534) 57 Carries out a risk-risk analysis (.739) 58 Considers the precautionary principle (606)
15 Impacts on non-profit sector	2.047% (50.948%)	69 Discusses whether regulation imposes costs on the non-profit sector (.845) 75 Discusses whether the non-profit sector benefits from regulation (.845)

Another way to validate the PCA is to control *ex post* whether the components cluster variables that are conceptually measuring the same construct. Indeed, we can confidently claim that all components do so.

THE WEIGHTED INDEXES

In a next step we created an index of IA quality. This is based on the sum of the components scores weighted by the share of variance explained by each component (see Table A4). The index relies on the 15 components identified above. The code reads as follows:

$\boldsymbol{Main\ Index\ score_i}$ = (0.193 * FAC1) + (0.099 * FAC2) + (0.078 * FAC3) + (0.072 * FAC4) + (0.063 * FAC5) + (0.061 * FAC6) + (0.058 * FAC7) + (0.054 * FAC8) + (0.053 * FAC9) + (0.048 * FAC10) + (0.047 * FAC11) + (0.046 * FAC12) + (0.043 * FAC13) + (0.042 * FAC14) + (0.04 * FAC15) - whereby FAC stems for the principal component score of each observation.

The sub-indexes are calculated in the same way. However, they use a subset of components for each index (see main text).

1) $\boldsymbol{CBA\ Index\ Score_i}$ = (0.66 * FAC1) + (0.34 * FAC2)

2) $\boldsymbol{Economic\ Analysis\ Index_i}$ = (0.327 * FAC4) + (0.264 * FAC7) + (0.214 * FAC11) + (0.195 * FAC13)

3) $\boldsymbol{Social\ Analysis\ Index_i}$ = (0.312 * FAC5) + (0.265 * FAC8) + (0.226 * FAC12) + (0.197 * FAC15)

4) $\boldsymbol{Environmental\ Analysis\ Index_i}$ = (0.537 * FAC6) + (0.463 * FAC9).

Furthermore, we rescaled the indexes to 0-1 values in order to make interpretation easier. Finally, we normalized the indexes to meet the distributional assumptions for parametric analyses. Specifically, they were transformed in squared root values, and a few extremes outliers were manually removed.

As a result, the main index and the sub-indexes 1, 2 and 3 are normally distributed according to the analysis of skewness and kurtosis. Sub-index 4, also after squared root transformation and the removal of extreme outliers still shows a kurtosis above 2, which questions its normality. However, because ANOVA is reasonably robust with regards to violations of normality and because of the high explanatory value of high scores, we decided not to manipulate the sub-index any further.

Having said this, in order to guarantee robustness and reliability we complemented the ANOVA on the environmental sub-index, used to test hypothesis 4, by its non-parametric equivalent (Kruskal-Wallis test). Its results are perfectly in line with those of the ANOVA (see below, Tables A22 and A23).

HYPOTHESIS TESTING

Table A7. Time trend. Descriptives

IA breadth and scope

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
2005	82	.55103258	.127049919	.014030316	.52311666	.57894851	.155774	.798699
2006	93	.55751112	.126746900	.013143041	.53140791	.58361433	.270538	.918428
2007	73	.59773202	.131070282	.015340616	.56715106	.62831298	.294008	.969195
2008	96	.58981285	.164477549	.016786920	.55648660	.62313910	.206696	.974378
2009	78	.61811502	.153792747	.017413600	.58344011	.65278992	.181735	.937545
2010	56	.58613444	.168007920	.022451003	.54114163	.63112726	.105714	.911107
2011	37	.65022297	.147072559	.024178580	.60118654	.69925940	.317637	.932817
Total	515	.58715423	.147378607	.006494277	.57439564	.59991282	.105714	.974378

Table A8. Time trend. ANOVA with polynomial contrast

IA breadth and scope

			Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)		.420	6	.070	3.306	.003
	Linear Term	Unweighted	.285	1	.285	13.474	.000
		Weighted	.303	1	.303	14.308	.000
		Deviation	.117	5	.023	1.106	.356
	Quadratic Term	Unweighted	.000	1	.000	.000	.998
		Weighted	.002	1	.002	.098	.754
		Deviation	.115	4	.029	1.357	.248
	Cubic Term	Unweighted	.027	1	.027	1.262	.262
		Weighted	.011	1	.011	.507	.477
		Deviation	.104	3	.035	1.641	.179
	4th-order	Unweighted	.058	1	.058	2.732	.099

	Term	Weighted	.044	1	.044	2.099	.148
		Deviation	.060	2	.030	1.412	.245
	5th-order Term	Unweighted	.006	1	.006	.306	.581
		Weighted	.005	1	.005	.250	.617
		Deviation	.054	1	.054	2.573	.109
Within Groups			10.745	508	.021		
Total			11.164	514			

Table A9. 2007 guidelines. Group statistics

	@2007_guidelines	N	Mean	Std. Deviation	Std. Error Mean
CBA_ind_sqrt	0	201	.4643995	.14921164	.01052458
	1	284	.5963037	.18410822	.01092481

Table A10. 2007 guidelines. Independent samples t-test on sub-index 1 on benefit-cost measures).

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Sub-index 1	Equal variances assumed	14.749	.000	-8.39	483	.000	-.1319	.01571833	-.16278889	-.10101940
	Equal variances not assumed			-8.69	474.16	.000	-.1319	.01516965	-.16171220	-.10209609

Pre-guidelines: 0
Post-guidelines: 1

Table A11. 2007 Guidelines. Group statistics

	2007_guidelines	N	Mean	Std. Deviation	Std. Error Mean
IA breadth and scope	0	205	.55914457	.123949603	.008657016
	1	288	.60755786	.160457343	.009455040

Table A12. 2007 Guidelines. Independent samples t-test on main index.

		Levene's Test for Equality of Variances						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
IA breadth and scope	Equal variances assumed	13.783	.000	-3.619	491	.000	-.04841	.01338

	Equal variances not assumed			-3.777	487.69	.000	-.04841	.01282

0 → Pre-guidelines
1 → Post-guidelines

Table A13. Establishment of the Regulatory Policy Committee. Group statistics

	RPC	N	Mean	Std. Deviation	Std. Error Mean
IA breadth and scope	0	403	.58104320	.144960954	.007221019
	1	105	.61245383	.158592576	.015477051

Table A14. Establishment of the Regulatory Policy Committee. Independent samples t-test on main index.

	Levene's Test for Equality of Variances						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference

IA breadth and scope	Equal variances assumed	.869	.352	-1.939	506	.053	-.03141	.01620
	Equal variances not assumed			-1.839	152.338	.068	-.03141	.01708

0 → Pre-RPC
1 → Post-RPC

Table A15 Sectoral specialization. Sub-index 2 on economic impacts. Descriptives (ANOVA non-significant)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
1	106	.65229691	.124417902	.012084535	.62833551	.67625831	.342853	.946768

2	79	.64781526	.128069441	.014408938	.61912927	.67650125	.008651	.936701
3	77	.63274283	.131707617	.015009476	.60284887	.66263679	.267613	.950446
4	255	.62397365	.138803014	.008692182	.60685573	.64109158	.124585	.999820
Total	517	.63472990	.133474464	.005870200	.62319747	.64626233	.008651	.999820

1 → Economic departments
2 → Environmental departments
3 → Social departments
4 → Residual departments

Table A16. Sectoral specialization. Sub-index 3 on social impacts. Descriptives.

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
1	106	.50240779	.173316986	.016834034	.46902901	.53578657	.027647	.999970

2	79	.49200827	.171822555	.019331548	.45352211	.53049442	.058171	.862675
3	77	.58493331	.136613664	.015568572	.55392581	.61594081	.265269	.928502
4	254	.54704141	.151263865	.009491139	.52834970	.56573311	.107586	.949011
Total	516	.53510128	.159802856	.007034928	.52128059	.54892197	.027647	.999970

Table A17. Sectoral specialization. Sub-index 3 on social impacts. ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.487	3	.162	6.569	.000
Within Groups	12.664	512	.025		

Total	13.152	515			

Table A18. Sectoral specialization. Sub-index 3 on social impacts. ANOVA multiple comparisons

	(I) Departmental specialization	(J) Departmental specialization	Mean Difference (I-J)	Std. Error	Sig.

Tukey HSD	1	2	.010399517	.023376080	.971
		3	-.082525522*	.023549384	.003
		4	-.044633622	.018185856	.069
	2	1	-.010399517	.023376080	.971
		3	-.092925039*	.025185797	.001
		4	-.055033139*	.020260213	.034
	3	1	.082525522*	.023549384	.003
		2	.092925039*	.025185797	.001
		4	.037891900	.020459927	.250
	4	1	.044633622	.018185856	.069
		2	.055033139*	.020260213	.034
		3	-.037891900	.020459927	.250
Games-Howell	1	2	.010399517	.025633834	.977
		3	-.082525522*	.022929569	.002
		4	-.044633622	.019325279	.100
	2	1	-.010399517	.025633834	.977
		3	-.092925039*	.024821144	.001
		4	-.055033139	.021535795	.057
	3	1	.082525522*	.022929569	.002
		2	.092925039*	.024821144	.001
		4	.037891900	.018233545	.165
	4	1	.044633622	.019325279	.100
		2	.055033139	.021535795	.057
		3	-.037891900	.018233545	.165

Table A19. Sectoral specialization. Sub-index 4 on environmental impacts. Descriptives.

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	105	.40183253	.088776352	.008663685	.38465212	.41901294	.249822	.773183
2	79	.51899383	.130724719	.014707680	.48971309	.54827457	.294870	.815718
3	77	.42322231	.089518414	.010201570	.40290412	.44354050	.143397	.672229
4	255	.44219374	.135810512	.008504784	.42544487	.45894262	.134219	1.000016
Total	516	.44290785	.125496906	.005524693	.43205414	.45376156	.134219	1.000016

Table A20. Sectoral specialization. Sub-index 4 on environmental impacts. ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.664	3	.221	15.229	.000
Within Groups	7.447	512	.015		
Total	8.111	515			

Table A21. Sectoral specialization. Sub-index 4 on environmental impacts. ANOVA multiple comparisons

	(I) Departmental specialization	(J) Departmental specialization	Mean Difference (I-J)	Std. Error	Sig.
Tukey HSD	1	2	-.117161301*	.017961491	.000
		3	-.021389784	.018094115	.638
		4	-.040361215*	.013983895	.021
	2	1	.117161301*	.017961491	.000
		3	.095771518*	.019312791	.000
		4	.076800086*	.015528563	.000
	3	1	.021389784	.018094115	.638
		2	-.095771518*	.019312791	.000
		4	-.018971432	.015681776	.621
	4	1	.040361215*	.013983895	.021
		2	-.076800086*	.015528563	.000
		3	.018971432	.015681776	.621
Games-Howell	1	2	-.117161301*	.017069718	.000
		3	-.021389784	.013384000	.383
		4	-.040361215*	.012140461	.005
	2	1	.117161301*	.017069718	.000
		3	.095771518*	.017899382	.000
		4	.076800086*	.016989620	.000
	3	1	.021389784	.013384000	.383
		2	-.095771518*	.017899382	.000
		4	-.018971432	.013281694	.483
	4	1	.040361215*	.012140461	.005
		2	-.076800086*	.016989620	.000
		3	.018971432	.013281694	.483

Table A22. Sectorial specialization. Sub-index 4 on environmental impacts.. Mean Ranks

	Departmental specialization	N	Mean Rank
Environmental sub-index	1	105	199.39
	2	79	350.30
	3	77	250.29
	4	255	256.88
	Total	516	

Table A23. Sectorial specialization. Sub-index 4 on environmental impacts. Non-parametric test, Kruskal-Wallis

	Environmental sub-index
Chi-Square	46.718
df	3
Asymp. Sig.	.000

Table A24. Core executive. Main index. Descriptives

IA breadth and scope

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
0	309	.60160195	.147229171	.008375575	.58512137	.61808254	.105714	.974378
1	132	.56048134	.148316839	.012909324	.53494362	.58601906	.181735	.937545
2	74	.57440387	.140244884	.016303137	.54191177	.60689598	.313544	.937113
Total	515	.58715423	.147378607	.006494277	.57439564	.59991282	.105714	.974378

0 → Non-core departments
1 → Core departments
2 → Super-core departments

Table A25. Core executive. ANOVA on main index

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.170	2	.085	3.969	.019
Within Groups	10.994	512	.021		
Total	11.164	514			

Table A26. Core executive. ANOVA multiple comparisons

	(I) Department type	(J) Department type	Mean Difference (I-J)	Std. Error	Sig.
Tukey HSD	0	1	.041120612*	.015236789	.020
		2	.027198081	.018964643	.324
	1	0	-.041120612*	.015236789	.020
		2	-.013922531	.021279977	.790
	2	0	-.027198081	.018964643	.324
		1	.013922531	.021279977	.790
Games-Howell	0	1	.041120612*	.015388336	.022
		2	.027198081	.018328735	.302
	1	0	-.041120612*	.015388336	.022
		2	-.013922531	.020795262	.782
	2	0	-.027198081	.018328735	.302
		1	.013922531	.020795262	.782

Table A27. Factorial ANOVA, interaction between year of publication and 2007 guidelines

Tests of Between-Subjects Effects

Dependent Variable: RIA analytical richness

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.458[a]	8	.057	2.679	.007	.042
Intercept	28.363	1	28.363	1326.861	.000	.733
Year_of_publication	.156	6	.026	1.220	.295	.015
@2007_guidelines	1.538E-5	1	1.538E-5	.001	.979	.000
Year_of_publication * @2007_guidelines	.009	1	.009	.419	.518	.001
Error	10.346	484	.021			
Total	180.924	493				
Corrected Total	10.804	492				

a. R Squared = .042 (Adjusted R Squared = .027)

Table A28. Factorial ANOVA, interaction between year of publication and establishment of the RPC

Tests of Between-Subjects Effects

Dependent Variable: RIA analytical richness

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.441[a]	7	.063	2.942	.005	.040
Intercept	103.426	1	103.426	4830.932	.000	.906
Year_of_publication	.359	6	.060	2.793	.011	.032
Dummy_RPC	.000	1	.000	.017	.895	.000
Year_of_publication * Dummy_RPC	.000	0	.	.	.	.000
Error	10.705	500	.021			
Total	186.506	508				
Corrected Total	11.145	507				

a. R Squared = .040 (Adjusted R Squared = .026)

Table A29. One-way ANOVA of main index using sectoral specialization as grouping variable. Descriptives.

RIA analytical richness

		N	Mean	Std. Deviation	Std. Error
1 -		104	.57570287	.143503793	.014071705
2		79	.61211071	.147911598	.016641355
3		77	.58843061	.146958206	.016747442
4		255	.58370755	.148872159	.009322736
Total		515	.58715423	.147378607	.006494277
Model	Fixed Effects			.147373066	.006494033
	Random Effects				.006556061

1 → Economic departments
2 → Environmental departments
3 → Social departments
4 → Residual departments

Table A30. ANOVA of main index using sectoral specialization as grouping variable.

RIA analytical richness

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.066	3	.022	1.013	.387
Within Groups	11.098	511	.022		
Total	11.164	514			

Table A31. The 2007 IA guidelines across departments

Typology of department	Number of pre-guidelines IAs	Number of post-guidelines IAs	Sub-index 1 mean pre-guidelines	Sub-index 1 mean post-guidelines
Economic departments	54	48	0.466	0.626*
Environmental departments	32	42	0.457	0.617*
Social departments	38	38	0.488	0.573*
Residual departments	77	156	0.454	0.587*

The asterisk indicates statistically significant differences (p<.05) in the t-test for equality of means. The improvement of sub-index 1 after the 2007 guidelines is homogeneous across departments.

---

i See Table A1 in our supplementary materials for a list of all UK departments including their abbreviated versions and Table A2 for an overview of our sample.

ii See Table A3 in in our supplementary materials for our scorecard.

iii See pages 7 to 16 in our supplementary materials for an in-depth discussion of how we conducted the PCA and, based on the PCA, constructed our dependent variables.

iv See Table A6 in our supplementary materials for further details on the principal components and their loading variables and coefficients.

v Tables A7 and A8 in our supplementary materials provide additional materials.

vi We wish to emphasize that we carried out two ANOVAs on the main index: one employing polynomial contrasts and another one using planned contrasts. Both ANOVAs are based on a comparison of the same marginal means. While the first ANOVA analyzes the trend of the main index over time including its shape and is used to test Hypothesis 1, the second ANOVA specifically contrasts the means of pre-electoral, electoral and post-electoral years and their statistically significant differences to test Hypothesis 4.

vii One-way ANOVA's contrast significance p=.028 – one-tailed.

viii Environmental departments perform significantly richer environmental analyses than economic (p<.01 – Games-Howell pot-hoc test; Cohen's d=1.045) and social departments (p<.01 – Games-Howell post-hoc test; Cohen's d=0.857).

[ix] According to the post-hoc multiple comparisons, social departments significantly outperform economic ($p<.01$ – Tukey HSD post-hoc test; Cohen's d=0.532) and environmental ($p<.01$ – Tukey HSD post-hoc test; Cohen's d=0.598) departments when it comes to the analysis of social impacts.

[x] In particular: Peripheral vs. Core departments ($p=.02$ – Tukey HSD post-hoc test; Cohen's d=.285).

[xi] Please note that the number of IAs used per departmental 'stack' may be slightly lower than the figures provided in the additional materials for the total number of IA per department. This is because at times IAs came with no precise publication date. We possess information on the publication year but do not know the day and month, making it impossible to integrate such an IA meaningfully into the departmental 'stack'.

[xii] Both DFES and DIUS had a strong focus on innovation, skills and the commercial exploitation of scientific achievements with a view to facilitating economic growth. Unsurprisingly, they were later collapsed with BERR to form BIS.