From chasing violations to managing risks: origins, challenges and evolutions in regulatory inspections.

Blanc, F.O.M.

5. Conclusion

As suggested in our introduction, we have attempted throughout this research to try and bridge the gap between "theory" and "practice" perspectives, i.e. to bring together perspectives issued from decades of research on enforcement methods, compliance drivers and risk regulation – with knowledge, experience and data coming from practitioners of regulation and regulatory reform.

We set out to investigate whether risk-based approaches to inspections, and more broadly what one could call "smarter" approaches to inspections and enforcement, appeared to live up to their promises – namely a "win-win" result of higher effectiveness (or at least constant effectiveness with reduced costs and burden), higher efficiency, and better economic results. In so doing, we also sought to validate the consideration of inspections as a specific and distinct object, to present some of the main variations in inspections practices, and to define more precisely what risk-based, "smarter" inspections consisted of – and what exactly was to be understood under the word "risk". We also intended to look at the issue of trust (between market actors, in the regulatory system), and the extent to which different inspection methods may influence the trust level. Finally, we intended to look into the question of outcomes – how they are defined and measured, how differently formulated goals may influence actions and results, and what measurement challenges exist in order to assess the impact of inspection practices.

5.1. Testing our hypotheses

If we now consider our three main hypotheses, there are grounds to consider them validated to a significant extent. First, that inspections are a distinct object with their own range of effects, distinct from that of the regulations they aim at implementing, appears strongly demonstrated. This stems both by the historical perspective, which shows the specificity of the institutional development of inspections, independently from regulations, but also by the comparative perspectives, where jurisdictions with similar legislation (e.g. EU countries on food safety) have very different situations and results linked at least in part to different inspections and enforcement methods.

Second, risk is an important instrument to define goals and indicators, allocate resources, select priorities, decide on proportionate enforcement measures – but it is also a word fraught with polysemy, and a phenomenon that gives rise to conflicting perceptions. The effective implementation of risk-based approaches in inspections and enforcement requires an adequate definition of risk (combining not only probability of hazard, but its potential magnitude), as well as data to ground planning and prioritization of actions. It also requires an enabling legal environment, which is sometimes problematic, and a political environment that likewise allows it, and does not impose a risk-averse approach. We have seen that, while there appears to be a strong justification for enabling risk-based discretion, forces that push towards risk-aversion are strong, and risk-assessment is sometimes difficult (uncertainty, lack of data), and/or conflicts with common perceptions.

Third, to the extent that data is available and can be relied upon, and leaving aside for a moment the question of causality and attribution, there exists some preliminary evidence that risk-based inspections practices produce better outcomes than other approaches, and that lower inspection numbers and more supportive enforcement practices do not lead to worse compliance levels, but rather (if done in the context of better targeting, "smarter" enforcement, more guidance etc.) to better safety levels (at least in some cases). The difficulty, however, is that data is far from sufficient to deliver strong evidence and very robust findings. Detailed, firm-level data is difficult to access, unavailable in many cases, frequently unreliable. Aggregate data is not always of high quality, and presents attribution issues. In addition, the question of economic impact is very difficult to properly investigate.

5.2. The case for risk-based inspections

Considering one of the most striking cases that we reviewed (occupational safety and health inspections and safety levels compared in Britain and Germany, and – with less detailed data – in France), it appears that the British HSE has long been conducting less inspections, putting more emphasis on risk-management and compliance-promotion, and achieved considerably better results than comparator countries. If comparing with Germany (where data is more detailed), Britain has been having consistently several times less frequent inspections and overall better outcomes (though Germany has been catching up). If comparing with France, safety in Britain is far higher, and so appears to be the overall trust in the system and trust between the different stakeholders (though this is based on non-quantitative evidence, the level of conflict around labour inspections in France is very high, with police protection regularly necessary, while effectiveness is clearly in question).

Of course, none of these findings is fully "robust" – cases were not randomly selected but picked because of data availability and knowledge about the different systems, labour inspections in France (and to some extent in Germany) check also other issues than OSH, we did not consider other safety indicators than fatal accidents (because they are known to be less reliable and/or, for occupational diseases, have a very strong time-lag), and the differences in performance may have a variety of causes. In addition, data on the economic impact of inspections is limited, so we had to use the number of inspections as a simple proxy (though Britain also has inspections that are far more "supportive" of businesses, guidance-oriented, and not only less frequent, thus there is a strong likelihood that they indeed create far less economic burden). Overall, the case is strong in favour of more risk-based and "smarter" inspections, but it is not without flaws in terms of data.

5.3. Challenges with data, and methods

What can be hoped for is that, this research having established with some level of evidence that the inspection issue is worthy of specific study, and that certain aspects deserve particular attention (targeting and enforcement models, safety outcomes etc.), there could be successive research undertakings focusing on specific inspection functions and jurisdictions, and attempting to use firm-level data to investigate correlations and possible causations in a more quantitative way. There would, however, be significant challenges in strict data-based investigations. First, firm-level data may be impossible to obtain (confidentiality or faulty information systems), and/or may require significant resources to conduct *ad hoc* surveys. Second, as we have attempted to show, there may be substantial problems with the quality of data, and with the interpretation of answers to surveys.

Overall, it is not certain that attempting to conduct such studies would be the best way to increase knowledge and understanding of the field. Were more governments to conduct representative business surveys assessing overall coverage and burden from inspections, as well as collecting some key "qualitative" data, and were more researchers to consolidated available data from inspectorates, our collective understanding of inspection activities would already greatly increase. If, in combination, more governments were to conduct surveys allowing to test the situation in terms of key public welfare outcomes (safety levels, prevalence of certain diseases, etc.), our knowledge of outcomes would be far greater than it is now. Even absent such additional resources and surveys, a deeper look at existing data sources (e.g. epidemiological) could allow to better assess the relative performance of different jurisdictions. Thus, rather than a focus on firm-level data research (which, as we have seen, seems to produce conflicting results, for a number of reasons), the field may be better served by more investigation of existing (consolidated) data at the level of entire jurisdictions and agencies, at least in a first phase. It is not clear, given the state of social sciences research and the issues

posed by data concerning complex social phenomena, that studies based on statistical analysis of correlations (and on testing causal models) would necessarily produce more useful results[691].

5.4. Limits and downsides of inspections

We think that this research at least shows the relevance of investigations of inspection practices and their effects, purported and real, challenging assumptions and "established wisdom". Indeed, the default assumption is still too often that more means better (and that stricter also means better). In reality, on the contrary, the effect of inspections is far from obvious. If we consider historical developments, the Netherlands set up a labour inspectorate only decades after France, Germany or Britain – it would be interesting to see whether the evolution of occupational-related deaths and injuries was markedly different or not. Indeed, available data suggests that improvements in such indicators took place regularly in the United States long before inspectorates were created. Likewise, major improvements occurred in food safety even at times when inspections were few and far between, conducted with methods and techniques that would not be considered adequate today. Improved technologies, better science, social movements, prevailing cultural norms etc. were all major drivers of these changes. How much of a role regulations and inspections play is still difficult to assess, but it may not be as much as some of their staunchest advocates claim.

The string of factory disasters in Bangladesh (and other countries of Southern and South-East Asia), which had as its most salient tragedy the *Rana Plaza* collapse[692] killing upwards of 1,100 people, illustrates the limits of inspections and enforcement. Even though the international outcry following it included many voices calling for more stringent inspections and enforcement, a closer investigation of the events leading to the building's collapse suggested this was not as easy (or as certain to produce results) as it sounded.[693] Indeed, inspectors had in fact responded to calls by workers that warned about the building being structurally unsafe, and ordered its closing, but their orders were simply disregarded by the owners. Thus, the disaster (and its causes) pointed to far more structural issues: weak rule of law (particularly for certain categories of powerful people), deep social inequality in terms of enforcement of legal rights, prevailing social norms among factory owners etc. Better targeted inspections and stronger powers for inspectors may be part of the solution, but they were (and are) far from certain to be sufficient.

---

[691] From our perspective, the current (disputed) "replication crisis" in social sciences reflects a number of issues, including less-than-optimal methods in a number of cases (bias against "null effect" reporting, excessively small samples, manipulation of samples to achieve statistical significance etc.), but is also likely to reflect the sheer complexity of the phenomena being studied, for which the variables measured may be inadequate proxies, and the difficulty to have reliable answers (or reliable interpretation of answers) on complex human behaviours and affects, when conducting surveys. We have already referred to the Open Science Collaboration (2015). Many other scholars have weighed in regarding replicability in psychology (and particularly social psychology). Pashler and Wagenmakers (2012) wrote that: "replicability problems will not be so easily overcome, as they reflect deep-seated human biases and well entrenched incentives that shape the behavior of individuals and institutions. Nevertheless, the problems are surely not insurmountable" (p. 529). Stroebe and Strack (2014) dispute the idea of a "crisis". For them, "because experiments are typically conducted with the aim of testing a theoretical hypothesis, the important question is not whether the original finding can be duplicated but whether it constituted a rigorous test of the postulated mechanism" (p. 62). In other words, if the original experiment was a poor test of the theory, replicating it will not give further "proof" of the theory – and not replicating it will not invalidate the theory either. Gilbert *et al.* (2016) have also challenged the Open Science Collaboration paper and findings, and write that they "did not take into account the multiple sources of error in their data, used a relatively low-powered design that demonstrably underestimates the true rate of replication, and permitted considerable infidelities that almost certainly biased their replication studies toward failure". Our tentative conclusion would be that, given the very high complexity of the interactions studied in the field of inspections, the very high chances for flaws in data or differences making replication impossible, the many ways in which design can have flaws that make the study inadequate to test the theory, this may simply not be the most productive approach, and a more modest approach using aggregate data and qualitative approaches may be more fruitful.

[692] See Wikipedia article: https://en.wikipedia.org/wiki/2013_Savar_building_collapse for a summary and sources

[693] Source: unpublished World Bank Group assessment in which the author was involved.

Increased inspections and "tough" enforcement may thus be neither strictly necessary to secure safety improvements, nor always effective at doing so either. They may in addition be harmful, if misguided because of a political "knee-jerk" reaction to a (real or purported) scandal. A recent illustration of what has been called in the Netherlands the "risk regulation reflex" can be seen in the nail salon "scandal" and its aftermath in New York City.[694] After the original reporting (by the New York Times) claimed rampant abuses, exploitation and unsafe working conditions, the City and the Governor responded with stepped up inspections and new regulations. These led to a significantly negative impact on businesses and their workers (closed salons, lost jobs), whereas further reporting a few months on showed the original report to have been nearly fully wrong (misrepresented facts, translation problems, exaggerations etc. led to its having little in common with reality). This is just an illustration of the fact that inspections are not a purely benign instrument – they can fail at their stated purpose, and they can also create actual harm. As such, their use needs to be carefully thought through.

5.5. Inspections and trust

To a significant extent, and this echoes one of our opening questions, an essential role of inspections and enforcement appears to create *trust*, and this may be the way in which their most positive effects are achieved – but even in this respect the risk of "overshooting" also exists. "Credible" inspections and enforcement are expected and requested by different groups of market participants and act for them as a guarantee that others will abide by the rules, and thus give them the appropriate assurance that they can engage in market interactions without excessive risk taking. In this respect, inspections and enforcement create a "floor" of confidence, and a risk-mitigation mechanism. They are also an important driver of compliance not only or even mostly through deterrence, but by reinforcing voluntary compliance: those who comply voluntarily have the confidence that the rules of the game are enforced and that those who are trying to undercut them through non-compliance do not have free rein. Credible inspections and enforcement reinforce the expectation of compliance, and thus increase trust among businesses (towards suppliers and buyers), workers (towards employers), consumers (towards sellers) etc. In an optimal situation, it reduces the reliance on litigation, reduces uncertainty, thus decreases costs and ends up having a positive economic and social impact.

There are however limits to the use of inspections to create trust. One of the earliest uses of regulations and inspections, as we have seen, has been to establish and enforce weights and measures, which in turn enable trade by simplifying transactions and reducing the risk for buyers[695]. This has early on started to evolve towards regulations and inspections covering the *quality* of goods, such as we have seen with the French *Inspection des Manufactures.* Such "quality inspections" are not, however, without problems in a market economy context. Defining quality is a very difficult proposition, at least when it goes beyond the definition of what a particular name should correspond to, and attempts to step into coercing sellers to provide only "good quality" wares. For instance, defining what can be called "olive oil" and what should only be called "vegetable oil", or what can be called "chocolate" and what should only be called "confectionary with cocoa" is a relatively narrow intervention, and can be grounded on clear norms, botanic and chemical definitions etc. By contrast, who is to determine (and how) whether garments, furniture or any other wares (or services) are of "adequate

[694] See successive reporting by: the New York Times, exposing the "scandal" (May 2015) http://www.nytimes.com/2015/05/10/nyregion/at-nail-salons-in-nyc-manicurists-are-underpaid-and-unprotected.html?_r=0 – the New York Review of Books, exposing factual errors (July 2015) http://www.nybooks.com/daily/2015/07/25/nail-salons-new-york-times-got-wrong/, and Reason followed by the New York Post fully deconstructing the original reporting, and showing the strong negative consequences it produced (October-November 2015) http://reason.com/blog/2015/10/27/new-york-times-nail-salon-unvarnished - http://reason.com/blog/2015/11/10/new-york-times-nail-salons-ron-kim - http://nypost.com/2015/11/03/the-new-york-times-refuses-to-face-facts-on-its-bogus-nail-salon-expose/

[695] This function is also important for other regulatory instruments such as licensing, certification or registration – cf. Kleiner 2006.

quality"? If such powers are given to inspectors, there is a real risk of abuse, decisions that are not grounded in clear requirements, excessive market restrictions, and plain and simple waste of resources.[696]

5.6. Closing remarks and questions

It is clear, thus, that this research's conclusion opens more questions than it gives answers. We hope to have shown that inspection institutions and practices, and their evolution and variations, are important objects that deserve to be researched and considered independently rather than simply amalgamated into "regulation" or "law enforcement". There is evidence that the effect of inspections is ambiguous and can be positive (on trust, on compliance, and to a somewhat unclear extent on safety and other public welfare outcomes), but also negative (on economic activity, employment, and also on trust and compliance if inspection practices are poor).

There is also evidence that risk-based approaches to inspections planning and implementation, and to enforcement decisions, seem to produce better results *both* in terms of safety, compliance and other targeted outcomes – but also in economic terms. The role of "risk" appears particularly central, with risk-averse reactions tending to produce poor regulatory outcomes, but an adequate understanding and management of risk seeming essential to balance different priorities and improve results. As the instrument allowing to assess priorities and to decide on what action to take and which resources to allocate, risk can be seen as the "currency of regulation"[697], i.e. the common unit that allows to "trade" between different priorities.

Specifically, risk-focused and risk-proportionate inspections and enforcement, combined with an approach that seeks to promote compliance by drawing on all its drivers rather than a narrow deterrence angle, can result in significantly improved public welfare. It also appears that such approaches work best when they combine:

- High level of professionalism and skills, not only in the narrow technical sense but encompassing a variety of competences (relations with businesses, risk management etc.);
- Balanced risk assessment methodology combining intrinsic risk with compliance history and confidence in management – supported by data and information technology to effectively target inspections and manage follow-up actions;
- Responsive and risk-proportionate enforcement, where the emphasis is on achieving increased compliance overall rather than symbolic punishment (even if the latter can be used in some cases);
- Active efforts to promote compliance through clear and reliable guidance, advice etc.;
- Transparency in methods, criteria, processes, decisions, combined with giving a real "voice" to duty holders, so as to maximize the sense of procedural justice and resulting legitimacy.

Such approaches can only be fully established in a context of agreement about goals, i.e. consensus that the primary objective of inspections is to achieve reduced risks and/or improved public welfare (in particular through an increase in compliance with specific regulations, but not only). In the absence of such consensus, e.g. for instance when the primary objective is seen as identifying and punishing violations, this combination of elements will be impossible (even if, for instance, targeting of visits is "risk based") and the results will unavoidably be different. Hodges (2015) has shown how there can be a very direct tension between achieving increased safety results and "punishing violations" (see e.g. the case of aviation safety pp. 326-329, where the

[696] See e.g. the case of Lithuania's Non-Food Products Inspectorate, which is forced by poorly-worded legislation and ingrained practices to conduct inspections upon consumer complaints, even when no legal norm has been infringed, i.e. pure "quality" inspections. These mostly relate to trivial issues, and have little results, but waste considerable resources for the inspectorate (and for businesses). See OECD (2015) pp. 132-134.

[697] This expression courtesy of Graham Russell, Director, UK Better Regulation Delivery Office

promotion of what is called as a "just culture" where reporting safety incidents is incentivized cannot, by definition, be done in a context of systematic punishment of violations).

Finally, there is a complex set of activities that are directly connected to inspections and enforcement tasks and institutions, but are distinct from inspections *stricto sensu*, such as provision of information and guidance. We have deliberately included some coverage of these in this research, as it is indispensable to properly assess inspectorates' activities and effectiveness, but a better word than "inspections and enforcement" may be needed to designate this field and this set of activities and institutions. In the United Kingdom, the expression "regulatory delivery"[698] has been adopted for this purpose – which also includes other regulatory instruments such as permits or licenses. Researching not only how inspections and enforcement are organized and conducted, but also in what broader "regulatory delivery" system they fit, and how different "regulatory delivery" models perform, also seem like important areas of future research.

Just as the field of inspections research as such, and particularly from a comparative context, is still relatively new (even though some very important works were written several decades ago, the "field" as such only emerged gradually over the last decade and a half[699]), policy interventions specifically targeting inspections are also relatively recent. Considering more thoroughly the different reform experiments[700], the different practices and their impacts, is clearly needed. We hope to have provided here some first elements that allow to show the relevance of such research, and to point at some directions it could take.

In addition, the question of methods to promote, support and verify compliance is important also for other areas of legislation – for instance the implementation of EU legislation by Member States[701], but also the "delivery" of legislation and rules applying to private citizens. Just as this research has drawn considerably from research originally focused on interactions between "law enforcement" (judiciary and police) and citizens, there is considerable potential for findings from studies of regulatory inspections and enforcement having relevance in other fields.

As a final word we hope that, in spite of its many limitations, this work will have shown the importance of challenging assumptions and beliefs when it comes to inspections and enforcement, and that there is a real possibility of "doing things better", in a way that benefits all stakeholders. In this respect, spreading what we *already* know is maybe as important as conducting further research to increase our knowledge, or test its limits.

---

[698] See the webpage of the UK's Better Regulation Delivery Office - https://www.gov.uk/government/organisations/better-regulation-delivery-office/about

[699] With "official recognition" internationally being achieved *inter alia* through the publication of the OECD's *Principles* (2014)

[700] As presented e.g. in Blanc 2012

[701] See Voermans 2015