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

Blanc, F.O.M.

4. Inspections and enforcement – a view from the practice

"What we need is to be told what is needed within the law and sound advice on how to complete certain things."

"No business gets everything right all the time. Where we fail despite our best efforts we would hope that we are not treated as if our non-compliance was deliberate."

LBRO, From the Business End of the Telescope *(2010)*

…reduction in scheduled inspections and other visible shifts away from enforcement activity (…) is sending a clear, calculated message to corporate criminals that (…), they will be even freer to kill and injure with impunity.

Steve TOMBS and David WHYTE, A Crisis of Enforcement *(2008)*

Street-level bureaucrats dominate political controversies over public services for (…)street-level bureaucrats have considerable impact on peoples' lives.

Michael LIPSKY, Street-Level Bureaucracy *(1980)*

While considering the many theoretical and scholarly accounts of inspections, enforcement, compliance and risk regulation, we have on occasion inserted references to current practices, and hinted at what could be learned from it. Before that, we also took the historical overview of inspections development up to the latest developments, and sketched out some of the main traits of different systems. We have not, however, investigated current practices as such, looking at the strengths and weaknesses of different regimes, and at whether we can assess their relative effectiveness (or to what extent we can do so).

In this chapter, we will attempt to give a "view from the practice". Not a comprehensive account, which would be an impossible task given our comparative focus, and already a very challenging task even if we only took one jurisdiction. Some have already proposed very rich accounts of inspections practices in one country (e.g. Mertens 2011). Others have offered detailed comparisons of one function, in at most a few jurisdictions (e.g. Tilindyte 2012). What we undertake here is both much less, and significantly more. Less, because we will have to limit ourselves to "snapshots", glimpses of different regimes and practices – no comprehensive account of any particular country or function. More, as we will try and have a broader reach, looking at a larger number of countries, and several functions.

First, we will sketch out several cases of inspections practices – both risk based (to varying extents and degrees), and clearly *not* risk based. We will try to show briefly both what these practices involve, and what their effects are. To the extent made possible by available data, we will try and compare some of the results and outcomes between different jurisdictions, to attempt to draw some lessons (however tentative) on the relative effects of contrasted approaches. Then, we will review and discuss some of the data issues, trying to

shed light on the limitations and constraints, as well as on possible ways forward – both regarding the use existing data, and the potential production of new data. Finally, we will conclude this chapter by trying to make sense of the different findings, considering what they can teach us about the problems inspections regimes face (or even contribute to create), the reform experiments that have been undertaken, and the results of both inspections practices and their transformations.

4.1. Case studies – views from the practice, comparative assessments

In this section, we will successively present three sets of case studies. The first will center on the British practice in inspections and enforcement, which can with some justice be presented as an exemplar of risk based approaches in general (Blanc 2012, Rothstein, Borraz and Huber 2013 etc.), and thus a good place to start to consider what these approaches look like in practice, and how they appear to perform. The second will consider the experience and problems of post-Soviet countries, as well as the contrasted reform trajectories and their apparent outcomes. Finally, the third will take somewhat shorter looks at the experience of several EU countries, including reforms and their limits.

We have attempted to keep the presentation of these case studies relatively short, because the overall scope of this research, and its length, were already significant. Significant work has gone into gathering and analysing the quantitative data that is presented in these case studies, and in particular in trying to make data comparable across different jurisdictions in spite of differences in sources, definitions etc. Considering that we have already exposed much of the historical and institutional background that is relevant to these case studies in chapter 2, however, we have sought to focus them on the essential aspects: description of methods (in particular in relation to risk, and to compliance management), presentation of available data, and discussion of findings.

A last word on the selection of these case studies. The intent of the research is to try and find an answer (even tentative) to the question of whether risk-based and "smart" inspection methods can yield "win-win" results, i.e. better outcomes in terms of public welfare and reduced burden from inspections for businesses. We have attempted to select cases that are meaningful to shed light on this question. First, the European cases (focusing primarily, but not exclusively, on OSH) build upon the historical research in chapter 2, with some variations in scope that correspond to the difference in focus (the intent is to look specifically at *diverging practices* to see whether they yield contrasting results), and to limitations in data availability. Second, the post-Soviet and post-Communist cases are useful because they cover a far larger set of countries and institutions (i.e. they add "scale" to the research), and they are a very strong illustration of non-risk-based approaches and their outcomes. In addition, significant attempts at reform (at least in some countries) mean that they offer an unrivalled opportunity to compare "before and after" (the move towards more risk-based inspections) in terms of both public welfare and administrative burden. Overall, we have a picture that combines focused research (OSH cases in Britain and Germany, with a less-detailed look at France and Italy as well), which shows very sharply contrasting results in both effectiveness and burden, with broader country-level pictures that provide a possibility of confirmation of the validity of findings in a broader context (larger set of jurisdictions and regulatory areas). Taken together, and because results appear strongly consistent, they offer a first element of response to the research question. Not a definitive one, of course, for which selected case studies could never be sufficient (and a definitive answer may well be out of reach in any case) – but at least a tentative one, suggesting that certain approaches yield consistently worse results than others, and that this should at least be ground to question them and look with greater attention to what makes more successful approaches work better.

a. Risk-based inspections in Great Britain – methods, practices, outcomes in OSH (and beyond)

Considering all the different aspects of inspections practices even in a single jurisdiction like Britain[483] would require a book unto itself. Rather, we will first focus on a function that is simultaneously one of the best studied already (e.g. Hawkins 2002, Tilindyte 2012), one where risk-based approaches have been developed for the longest time, and where data is relatively easy to access: occupational safety and health (OSH). To try and assess the effectiveness of OSH approaches in Britain, we will attempt a comparison of outcomes data with Germany[484], and see what lessons we can draw from it. In the conclusion to this comparison, we will also include highlights of some key aspects of risk-based approaches in food safety inspections, trying to show in what ways it most strongly differs from (to use a simple moniker) more "traditional" approaches in other EU countries. In so doing, we will also indicate when such approaches are used *beyond* food safety and OSH.

i. Context and evolutions in the past two decades – the consolidation of "risk based approaches"

We have sketched out, in the first chapter, the birth and evolution of the OSH regulatory system in Britain, and how it resulted in a dual structure – the Health and Safety Executive (HSE) responsible for (broadly speaking) "high risk" categories (at least as they were traditionally defined) and "major hazards" (in particular those covered by EU directives), and local authorities (LAs) dealing with "lower and medium risk" categories (again, based on what was understood as such several decades ago). In recent years, spurred by successive reports (e.g. Löfstedt 2011), the HSE has also taken a stronger role of guidance and coordination of methods, even though it does not have direct authority over LAs officers. Even prior to this recent trend of growing HSE involvement in guidance for LAs, a level of consistency was already ensured – both by the HSE's role in issuing clarifications and guidance on how to comply with regulatory objectives (which were used by businesses, and inspectors, regardless of which supervisory authority they reported to), and by the common background shared by many inspectors. Indeed, most LAs inspectors and a significant share of the HSE's[485] are "Environmental Health Officers" (EHOs), certified by the Chartered Institute for Environmental Health (CIEH)[486].

The existence of a specific profession of EHOs, with a broad perspective on environmental health risks rather than a narrow technical focus, as well as the large share of this profession employed in the regulatory sphere, are in themselves a specificity and have no correspondence in most other countries. Elsewhere, by contrast, such inspectors would stem from distinct technical fields, and have a background that is not linked to

---

[483] Because of the complex structure of the United Kingdom, regulatory structures are distinct in its different constituent parts – England, Wales, Scotland and Northern Ireland. In practice, England and Wales share national-level agencies, and also have similar structures at the local level. Scotland shares *some* of these with England and Wales, and the similarities are sufficient to speak of "British" practices (see details further in the text to justify this). While Northern Ireland's regulatory bodies tend to use similar approaches, their structure is sufficiently different to make it difficult to cover "UK practices" as a whole, hence our choice of "Britain" as the jurisdiction being considered.

[484] We will also make a partial comparison with France, but data on this country is incompletely available and does not allow for a full comparison. The choice of Germany is primarily due to Tilindyte's 2012 work comparing OSH in Britain and Germany, to which we are much indebted, as well as to the full availability of data, and the relative "proximity" of the two countries – but also to some other (real or imaginary) characteristics of Germany, as we will discuss further.

[485] Consistant with its missions, which focus far more on manufacturing industry, in particular heavy industry, and major hazards, the HSE also has a significant number of enginees and other technical specialists on staff, whereas LAs rely primarily on Environmental Health Officers.

[486] For an overview of its history and role, see CIEH's website at: http://www.cieh.org/about_us/history.html - for Scotland, CIEH's role is assumed by the Royal Environmental Health Institute of Scotland (REHIS) – see: http://www.rehis.com/about/about-rehis

regulatory issues. This specificity clearly has major significance in ensuring greater consistency, and in giving inspectors a sense of the existence of an "inspection and enforcement" field as such that cannot be reduced to the technical issues covered. Because food safety inspectors (at least those dealing with the processing and retail stages[487]) and environmental inspectors in the UK also generally are EHOs, this unique training and qualification model also ensures some consistency of approach and sharing of views beyond the OSH field.

Ensuring further coherence, and culminating in the recent evolution giving a stronger guidance role to the HSE, there has been a succession of policy steps to achieve greater consistency in enforcement. In 1997, an "Enforcement Concordat" was launched by the Local Government Association (LGA) and the Government (Secretary of state for the Environment, Transport and the Regions). This concordat was proposed for adoption by LAs, and established a number of "principles of good enforcement policy" (Davey 2011, pp. 263-264): clear *standards* for the level of service and performance, *openness* including information/advice and consultations, *helpfulness* based on a "prevention is better than cure" approach, effective and timely *complaints* procedures, *proportionality* of both requirements and enforcement actions and *consistency* (including with other enforcement bodies) to balance necessary inspector discretion. We have already reviewed above the HSE's Enforcement Policy Statement (see 3.3.c), and it is clear that there is a strong alignment between the two documents. In spite of institutional fragmentation (only one national-level body for OSH, but 433 local councils in the UK, including 407 in Britain, as of 2009[488]), there is thus a significant level of shared principles across all of Britain (and the UK).

This was further reinforced by successive developments from 2005 onwards. In 2005 was released the report of the review led by Philip Hampton: *Reducing administrative burdens: effective inspection and enforcement* ("*Hampton Review"*). OSH was of course among the major regulatory functions reviewed, both at the national and local level (see list of main regulatory functions pp. 14 and 17). This report was very influential and led to a series of Government initiatives. As a conclusion of the review of the existing situation and challenges, it proposed a series of principles, including: the need to use *comprehensive risk assessment* to target resources, *accountability* for efficiency and effectiveness combined with *independence* in operational decisions, regulations should be *clear* and based on *consultation*, inspections should be *justified*, information should only be requested *once* and when strictly *needed*, sanctions should *proportionate* but also *meaningful* and *prompt* for repeat offenders. Regulators should "provide *authoritative*, *accessible advice* easily and cheaply". In addition, the *Review* recommends consideration of enforcement issues from the policy drafting stage, avoiding the creation of new regulators when not needed, reviewing their "size and scope" compared to their missions, and recognition by regulators that "a key element of their activity will be to allow, or even encourage, economic progress and only to intervene when there is a clear case for protection" (p. 7). The *Hampton Review* has often been seen as being the starting point of a drive for more risk-based inspections in the UK. As we have seen, while it certainly put an increased emphasis on risk-based targeting and proportionality (and on the burden-reduction and economic growth angle) it built on a much longer tradition, and reiterated many of the principles already present in the 1998 concordat, for instance.

The *Hampton Review* resulted in a series of policy initiatives, which included the adoption of the 2008 *Regulators' Compliance Code* (see Davey 2011 p. 264). It was followed up by another review on sanctions and enforcement (*Macrory Review* 2008), which in turn led to the Regulatory Enforcement and Sanctions Act 2008, itself resulting in the creation of the Local Better Regulation Office (LBRO), in particular with a view to administer the new *Primary Authority* scheme. LBRO's mandate was to introduce more consistency in

[487] See our first chapter on the historical factors that led to the unusually weak role of veterinarians in food safety in the UK. Even though veterinarians are in charge of the Meat Hygiene Service (under the Food Standards Agency), the bulk of inspections remains handled by LAs, and primarily by EHOs.
[488] See LBRO 2009 (p. 25)

inspections and enforcement, in a spirit of risk-based regulation, reduction of administrative burden and compliance promotion. Further developments then included the transformation of LBRO into the Better Regulation Delivery Office (BRDO) in 2012, with a mandate now including national regulators in addition to local ones, and the adoption of the 2014 Regulators Code, superseding the 2008 one. These successive developments further consolidated the emphasis on the "Hampton Principles": risk-based targeting, proportionality, guidance and advice, attention to economic growth etc.

Several specific reviews were conducted of the Health and Safety regulatory system – the first in 2008 (done by the Better Regulation Executive) and a second one (after the change of majority) in 2011 (by R. Löfstedt). The 2008 review showed, among other issues, that the split of responsibilities between HSE and LAs, based on a rigid legislative formula[489], resulted in a misallocation of resources from a risk perspective that has been nicknamed the "twin peaks" problem. Both HSE and LAs targeted their resources based on a "risk pyramid", but the top of the LAs risk pyramid tended to be "lower" from a risk perspective than the base of the HSE one (see pp. 59-66). Thus, some premises with actually higher risk tended not to be inspected at all (or insufficiently) because they were in the HSE's remit, and some premises with in fact lower (though not inconsequential) risk were being inspected far more intensively. This showed the limits of the rigid allocation of responsibilities (which has upside from a clarity perspective and avoids overlaps, but cannot accommodate a full risk-based approach). As for the 2011 review, it emphasized the inconsistencies created by the division of responsibilities and the large number of LAs involved (pp. 78-83) – and as a result recommended giving HSE a much stronger coordinating role. It also, again, highlighted the "twin peaks" problem and called for a much narrower focus on high-risk premises (pp. 5 and 82-83).

The *Löfstedt Review* recommendations were translated into the 2013 *National Local Authority Enforcement Code*[490] issued by the HSE and applicable to all LAs enforcement activities relating to health and safety. Its purpose was summarized by the Government as such: "local authorities are being banned from unnecessary health and safety inspections" and "will instead target proactive council inspections on higher risk activities in specified sectors or when there is intelligence of workplaces putting employees or the public at risk"[491]. The *Code* sets out to impose on all LAs a series of principles and rules, all emphasizing the need to focus resources on higher risks, and to use methods proportional to the risk level[492]. In particular, "proactive inspection" should be used "only for premises with higher risks or where intelligence suggests that risks are not being effectively managed" (p. 2). The *Code* emphasizes the importance of "choosing the most appropriate way of influencing risk creators and by targeting their interventions, including inspection, investigation and enforcement activity, on those businesses and sectors that represent a higher level of risk to the health and safety of workers and the public" (p. 4). It also lists a number of roles for the HSE in relation to LAs, in particular making its advice and guidance "authoritative", and giving it a far stronger role of strategy definition, priorities setting, support and guidance for LAs (p. 5). The *Code* specifically requires LAs to have "risk-based intervention plans" and use "proactive inspections" only in "sectors specified by HSE" or where "intelligence" suggests problems with risk management (pp. 6-7). It also emphasizes the importance of proportionality, and instructs LAs to follow the HSE's *Enforcement Management Model,* as well as HSE's criteria for dealing with complaints on a risk basis (pp. 7-8). It also underlines the importance of the Primary Authority scheme as a way to provide more guidance and consistency, and more risk focus – as well as the need to develop LAs inspectors competences in line with

[489] The precise split is set out in the 1998 Enforcing Authority (Health & Safety) Regulations but "remains largely as it was in the early 1960s" (BRE 2008 p. 57).
[490] See the HSE website for an introduction and the full text: http://www.hse.gov.uk/lau/la-enforcement-code.htm.
[491] Taken from the Government website's summary page on the Code – see: https://www.gov.uk/government/news/new-code-curbs-unnecessary-council-safety-checks
[492] The *way* in which this is done raises a number of problems, which we will discuss further. For now, we will just note that the *Code* explicitly strengthens the importance of risk-based approaches.

the "competency approach" (pp. 9 and 11). In conclusion, following the adoption of the *Code,* a relatively unified model of risk-based inspections applies, at least in theory, to all health and safety inspections in Britain.

Decreases in staffing and number of inspections

For a variety of reasons, which it would go far beyond our scope to explore, "health and safety" (as it is generally referred to in the UK, since the regulatory remit is broader than only *occupational* risks) has been particularly targeted by efforts to reduce regulatory costs (for the state) and burdens (for businesses and citizens). This has led to a significant decrease in staffing levels, and in the number of inspections, at least for the HSE (it is far more difficult to assess such trends for LAs). One possible explanation for this decline is political (pressure from employers to reduce workers' protection and the associated costs), but (apart from ideological motives, which are of course possible) this would not explain why the HSE in particular has been under constant pressure to reduce inspections, more (or so it appears) than other regulatory areas.

First, there are only few major national regulators conducting a significant number of inspections in Britain – mostly, the HSE, the Environment Agency, and Her Majesty's Revenue and Customs[493]. Thus, HSE inspections "stand out" far more than, say, food safety ones, which are conducted essentially by LAs (except the meat slaughter ones, which are EU-mandated and thus "immune" to "burden reduction"). Second, there may be (but this would have to be investigated further) a difference in perceptions by the public. While food safety requirements and inspections impose clear costs on businesses, they are rarely visible by the public, and do not *seem to* result in a reduction in food availability (even though, in fact, they do). By contrast, there is a significant proportion of OSH requirements that appear to have a high "annoyance" factor, and are seen as limiting or preventing activities that used to be possible (more) freely. This public perception issue appears serious enough that HSE has a dedicated set of activities to address it: "busting the health and safety myths[494]". There are additional issues at play. Perceptions of health and safety burden by businesses and the general public (and the resulting "health and safety myths") are very probably driven as much (or possibly far more) by the activities and recommendations of private consultants, insurers and other actors than by HSE inspectors, who are far more likely to come up with "realistic and reasonable" solutions, at least according to many practitioners. In the end, however, the public does not differentiate where messages come from, and end up complaining about "health and safety" generally – which the Government then tends to react to as if it were an indictment of regulatory bodies in charge, HSE in particular[495]. In addition, the position of the HSE as a non-departmental body under the Department for Work and Pensions likely plays a role as well. Indeed, the HSE's functions bear very little relation to the Department's primary focus (delivering benefits). Thus, in a context of budget restrictions, cutting staff and resources to the HSE is likely to be frequently the path of least resistance, as it will not be seen as threatening the Department's overall performance.

Whatever the causes, and regardless of how these trends may be interpreted, there has been a significant decline in HSE staffing and budget, and in the number of HSE inspections. We saw in the first chapter that, between 2002 and 2014, overall HSE staff went down by at least a quarter[496]. The decline in the number of

---

[493] Tax inspections may also be on a downward trend but, given their link to revenue, they tend to be an area that Governments are less keen to make savings upon. Environmental inspections in England and Wales are also generally down, but primarily linked to changes in the permitting system, so in a different context from the HSE's.

[494] See on the HSE website: http://www.hse.gov.uk/myth/.

[495] See Dunlop 2014. Summary here: http://www.exeter.ac.uk/news/featurednews/title_427731_en.html

[496] Taking into account the spin-off of the Office for Nuclear Regulation (ONR) and thus not counting the nuclear safety staff in 2002 – the exact percentage is difficult to calculate given that the 2002 annual report provides nuclear safety staff separately, but only for operations, whereas the spin-off also included management staff for the ONR. Comparing "all of HSE" in 2002 with "HSE and ONR" in 2014, the decline is 24%. Taking only HSE without nuclear directorate in 2002, and HSE without ONR in 2015, the decline is 29%. Sources: successive HSE annual reports – this ans subsequent data from the HSE annual reports from 2001-2002 till 2014-2015 - see: http://www.hse.gov.uk/aboutus/reports/index.htm

inspectors is even sharper, particularly in recent years. They were 1,625 in 2002, and only 1,038 as of 31 March 2015[497] - a decrease of 36%[498].

While data on the number of staff is readily available, numbers of inspections have stopped being publicized by HSE nearly a decade ago, and have to be found in secondary sources (which obtained them e.g. through Freedom of Information requests) or "reconstructed" from different HSE publications. In the 2001-2002 financial year, there were 75,237 inspections (out of which 65,000 by the Field Operations Directorate). In 2002-2003, it went up to 84,234 (out of which 74,112). From there, it was a constant decline. In 2004-2005, Field Operations Directorate (FOD) inspections were down to 55,195[499]. In 2006-2007, the same number was only 41,496[500] - 44% less than in 2002-2003. The decline continued in later years. In 2011-2012, FOD undertook 21,603 proactive inspections. To these should be added around 4,000 reactive inspections (*HSE Annual Report 2011-2012,* p. 23) – 3,957 to be precise, even though not all of these were necessarily handled by FOD (but certainly the bulk of them were). HSE also "followed up circa 10,400 conventional health and safety complaints" (*ibid.*) – but not all of these "follow ups" were inspections.

The critics of this evolution, such as Tombs and Whyte or *Hazards Magazine* certainly have a point that the total inspections data is being dissimulated on purpose, making it difficult to track evolutions precisely. The question is, of course, whether HSE management (and the Government) are right that the number of inspections is simply misleading and irrelevant. Before moving to considering outcomes, however, let us try and look at the latest data to establish the current picture. HSE reported completing 5,004 investigations including a total of more than "over 3,260 incidents" meeting HSE criteria. The report also noted the "completion within agreed timescales" of 70% of complaints meeting HSE criteria and due to be followed up (*HSE Annual Report 2014-2015,* pp. 27-28) – meaning approximately 9,870 complaints followed up in a timely way. Though the report does not indicate it, only a minority of these follow ups involve inspections, and these are most probably counted within the 5,004 investigations noted above. A total of 20,200 proactive inspections were conducted (p. 16)[501], and "over 1,000 major hazard operators" were inspected (p. 31). Thus, while a grand total is not available anywhere, it can safety be assumed that the total number of all inspections conducted in 2014-2015 by the HSE did not exceed 30,000. Because each source of data and each year ends up giving us slightly different perimeters, we cannot do an exact estimate of the decrease for each year, but between 2002 and 2015 it was nearly two thirds (64%). This may somewhat over-estimate the decline, because 2002 itself came after several years of *increase* in resources and inspections at the beginning of the New Labour period. Still, it is quite a major decrease, and would still amount to a halving of inspections compared to 20 years earlier.

HSE inspections are, however, only a part of the total OSH inspections in Britain. While the HSE is responsible for a bit under half of total business premises, LAs are responsible for slightly more than half[502]. In terms of

---

[497] Note that the annual report 2014-2015 differentiates between "frontline staff (total)", "frontline inspectors" and "inspectors in functions othe than frontline". The overall picture is similar whether one counts "frontline staff (total)" (1047) or "all inspectors" (1038).

[498] The majority of these inspectors are in the FOD, but a large number are in the different specialized departments within HSE. The number of Full Time Equivalent (FTE) inspectors may in fact be lower, though the difference is sharper for LAs (see BRE 2008 p. 58).

[499] *Hazards Magazine,* number 94, April/June 2006 – article available at http://www.hazards.org/commissionimpossible/hse.htm (quoting HSE data obtained through Freedom of Information appeal).

[500] Quoted in Tilindyte 2012 p. 117 – see also Tombs and Whyte 2010.

[501] Thus implementing a target of the Coalition Government to see HSE "reduce its proactive inspections by one third (around 11,000 inspections per year)" (*Good Health and Safety, Good for Everyone - The next steps in the Government's plans for reform of the health and safety system in Britain* - 21 March 2011 – paper by the Department for Work and Pensions, available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/66745/good-health-and-safety.pdf – p. 9)

[502] See *Good Health and Safety, Good for Everyone* p. 10, suggesting a 50/50 split, and BRE 2008 p. 59 indicating a 55/45 split LAs/HSE. HSE's own data gives a slightly different picture: "HSE has responsibility for securing compliance in over 740 000 establishments and

employees, the LAs-supervised businesses represent close to 50% (reflecting the fact that, on average, LA-supervised establishments are slightly smaller)[503]. As of 2008, LAs also had more inspectors than the HSE's FOD – a total of 1,100 FTE working on health and safety issues (BRE 2008 – 3,320 inspectors in total, but covering a range of issues). Parallel to the decrease in HSE staffing, this number has decreased to 2,390 in total and 800 FTE in 2013-2014 – still more than the HSE's FOD[504]. The rise in importance of LAs in health and safety inspections largely reflects deep transformations in the economy, whereby "since the late 1970s, levels of employment and numbers of businesses have grown in the sectors of the economy that are inspected by local authorities" (BRE 2008, p. 58). LAs overall conduct far more inspections of health and safety issues than the HSE – even though a number of these are not "pure" health and safety but cover other issues as well, taking opportunity of the fact that EHOs have a competence that also cover for instance food safety issues[505]. Still, even with this noted, LA inspections are far more numerous than HSE ones (though they may often be shorter). In 2006-2007, HSE conducted "around 36,000" preventive inspections and LAs "around 121,000" (BRE 2008 p. 58 – noting also that HSE inspectors spent "significantly more time" per visit). The decline, however, has also been significant in the past few years for LAs inspections – in line with the Coalition government's objective to reduce inspections by LAs by "at least a third" (out of 196,000 in total, proactive and reactive, as of 2009-2010)[506]. In 2013-2014, after several years of reduction, the total number of LAs inspections for health and safety was down to 86,900[507].

Interestingly, the implementation of these new guidelines and priorities seems to have been done with a certain degree of confusion in objectives and methods – or at least in statistics. The *Enforcement Code* for LAs called for a reduction in proactive inspections, and ensuring they were focused only on high risk premises (or on premises where 'intelligence' suggested a high probability of non-compliance). The Government plans called for a reduction of ⅓ of inspections in total. In practice, LAs reported[508] cuts so radical in proactive inspections that these all but disappeared: while they made up 60% of all inspections in 2009-2010, they barely reached 8% of the total in 2013-2014. Reactive inspections (including follow-up inspections where problems were identified) increased, from slightly over 30% to around 45%. The gap was bridged by the "other" category, which rose from 9% to around 47%. Considering such figures, it is clear that the new policy has generated a high level of confusion, led to an increase in reactivity vs. proactivity (which is generally *not* a good thing in an approach aimed at *preventing risks*, since reactive inspections come, by definition, nearly always "too late") – and also led to a collapse in the meaningfulness of reporting categories, since the "other" group (supposed to be used for rare cases that did not quite fit one of the main categories) now makes up nearly half of the total. This move may, judging by the experience of many other countries, reflect a defensive move by inspectors and LAs disagreeing with the new policy, and deciding to keep to the approach they consider correct but game the system by avoiding to report their visits as "proactive" and selecting "other"

---

local authorities enforce the HSW Act in around 1 194 000 establishments" (*HSE Annual Report 2001-2002*), but successive reports have concluded that HSE's data was not fully accurate on this point.

[503] BRE 2008 p. 59.

[504] See *Data Collection – analysis of LAE1 2013/14 data from Local Authorities, Paper Number: H17/01*, Paper prepared for the HSE / Local Authorities Enforcement Liaison Committee – available at: http://www.hse.gov.uk/aboutus/meetings/committees/hela/

[505] See BRE 2008 p. 64 indicating that the number of inspections to catering premises reported included 60% of joint food inspections. Since 2011, such joint inspections have become the norm, see HSE, FSA and Local Government Regulation joint note: http://www.hse.gov.uk/lau/pdfs/combining-health-safety-and-food-safety-inspections.pdf

[506] See *Good Health and Safety, Good for Everyone* p. 10

[507] See *Data Collection – analysis of LAE1 2013/14 data from Local Authorities, Paper Number: H17/01*, Paper prepared for the HSE / Local Authorities Enforcement Liaison Committee – available at: http://www.hse.gov.uk/aboutus/meetings/committees/hela/ - also see latest detailed statistics on LA inspections at: http://www.hse.gov.uk/lau/enforcement-lae1-returns.htm - the detailed data only covered approx. 87% of local authorities. 86,900 is an extrapolation done by the paper's author (methodology unknown). Our own estimated extrapolation is around 85,000 (based on the number of enterprises or the population of the non-reporting LAs). The slight difference may come from the author using the employed population or another variable for extrapolation.

[508] See *ibid.*

instead. A more positive development appears to be the increasing focus on high risk objects – "inspection of higher risk premises has remained fairly constant but inspections of lower risk premises has more than halved since the introduction of the Code"[509]. This picture of increased focus is, however, unequal: "fewer than 10% of LAs account for over 77% of the lower risk inspections (B2/C) reported"[510]. Finally, it appears that many of the "other visits" are of advisory nature (a category of visits that the *Code* allows to continue without restrictions) – but in fact this again suggests that many of the former "proactive visits" (which were, to a large extent, aimed at advice and guidance to prevent risks) were mostly "renamed"[511]. While most of the evolutions more-or-less matched the guidelines issued by the Government, concern was expressed that, in nearly 20% of LAs, no inspections at all were undertaken[512].

Consequences of changes – disputed assessments

As Hawkins (2002) has shown in details, the use of prosecution has already long been a "last resort" for HSE inspectors – and we have indicated in the first chapter that this went back to the 19th century. Prosecutions, as a result, have always been rare, even though *in principle* health and safety violations are to a large extent "criminalized", i.e. *can be* subject to prosecution and (in case of conviction) criminal penalties. As a far more frequent alternative to prosecution, and in cases judged serious enough for simple advice to be insufficient, inspectors (HSE and LAs) can issue improvement notices (mandating the resolution of a given violation in a set time period) and prohibition notices (adding to this obligation the prohibition to use a given equipment, part of facility, entire establishment etc.). The latter, in particular, are quite powerful tools, as the economic damage imposed can be considerable. Tilindyte (2012) has concluded that the strength and flexibility of notices was such that HSE inspectors and management showed little interest to use the new Regulatory Enforcement and Sanctions Act 2008, and the possibility it offered to introduce new administrative penalties in addition to existing options of criminal prosecution and notices (see pp. 249-250 and 257-266). Many critics of the changes in health and safety enforcement in the past 15 years have spoken of a complete collapse in enforcement – *Hazards* magazine speaking of a "neutered watchdog"[513].

The data on *enforcement*, as distinct from that on *inspections*, is however far less clear than critics of successive reforms make it out to be. Tombs and Whyte (2008) show a significant decrease in HSE prosecutions from 2000-2001 to 2006-2007, but the picture is less clear on enforcement notices. These remained roughly constant for LAs, and the decline observed in HSE notices from 2002 to 2006 came after a significant increase from 1996 to 2002, and thus could be interpreted as a return to longer-term trends (p. 9). As Tilindyte (2012) shows, 2005-2006 was in fact (for whatever reasons) a low point in the number of notices, and these increased significantly afterwards, returning in 2009-2010 to a level that, while lower than the early years of Labour Government, was higher than in the years of Conservative Government in the first half of the 1990s (p. 140).

---

[509] See *ibid.* p. 2.

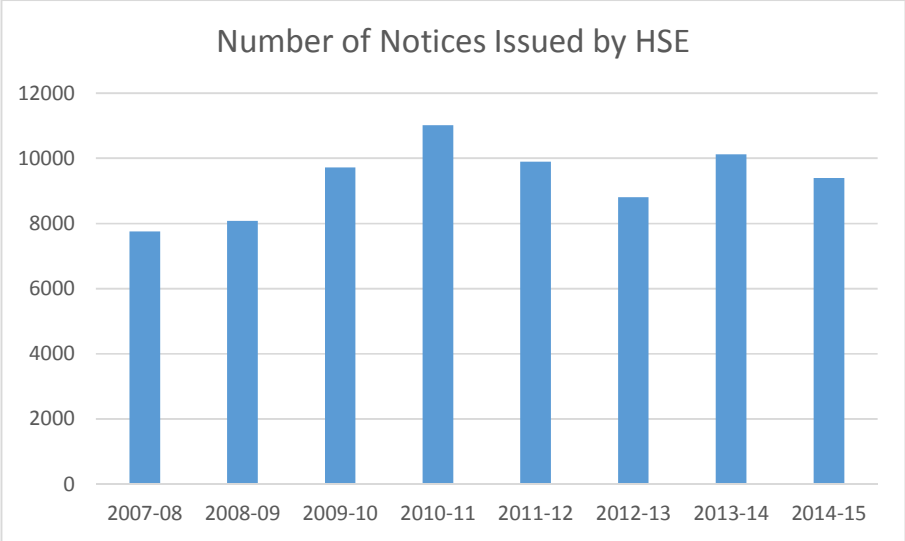
[510] See *ibid.* The note did not include a discussion of the *factors* that could lead to this difference in practices, and attempting to investigate it would have required considerable time. Experience from other countries would suggest that differences in "inspectors culture" and management vision could help explain it, as well of course as differences in the businesses themselves.

[511] See *ibid.* pp. 2-3. Advisory visits cannot result in sanctions in case of violations – but, as is well known from studies such as Hawkins 2002, BRE 2008 etc., health and safety inspections in Britain very rarely result in sanctions anyway. Hence, the "transformation" of many "proactive inspections" into "advisory visits" is mostly a case of change of label, rather than of substance, suggesting again the well known fact that excessively rigid quantitative targets, imposed without consideration to practice, often result in "gaming the system".

[512] See *ibid.* Once more, difficult to say how much it could have to do with local business circumstances (low-risk premises only) or with local priorities. In particular, even though LAs regulatory services are organized on a professional basis, they report to the local councils, which may have very different political views. One could imagine that Labour councils would support more health and safety inspections compared to Conservative ones (while having only limited ability to diverge from national trends due to legislative and budget constraints).

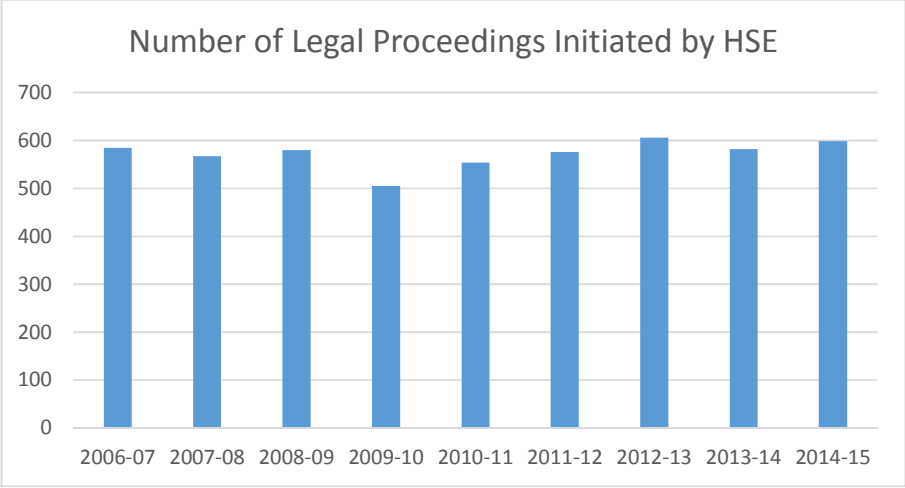
[513] *Hazards* issue 111, July-September 2010 – available at: http://www.hazards.org/votetodie/neutered.htm

Notices have continued to increase and decrease irrespective of the Government in place (as the chart below illustrates), and of reform trends – suggesting strongly that HSE inspectors are, in fact, quite independent in their assessments.



(Sources: *HSE Annual Reports 2011-12 and 2014-15* – data for 2014-15 is provisional)

As for prosecutions, as noted by Tilindyte (2012), there was first a noted increase in 1996-2000, followed by a decrease until 2006 (p. 142). Since then, the annual number of prosecutions has been relatively stable, between 500 and 600 a year, with fluctuations again seemingly not connected in any way with political changes (see chart below)[514].



(Sources: *HSE Annual Reports 2011-12, 2013-14 and 2014-15* – data for 2014-15 is provisional)

While HSE has always sought to focus its prosecutions on cases where it assessed that the chances of successful conviction were maximal (see Hawkins 2002, Tilindyte 2012), this has become an official

[514] Note: different sources and works use "informations" (one for each alleged offence) and "prosecutions" (one for each establishment) and thus give very different totals. The trends remain: notable decrease until 2006, stable since then – at a level lower than in the early 1990s.

performance target and the conviction rate is now regularly reported in annual reports. It has been around 95% for the past few years. While this speaks to an efficient use of resources from a narrow perspective, putting this as a performance target may twist incentives in a way that discourages inspectors to attempt prosecutions that may fit many criteria of relevance, but would be somewhat more difficult to conclude successfully. As Tilindyte indicates, there are "several set of explanations put forward" by HSE management for the decrease in prosecutions, in particular the "rebalancing of resources towards more advice and guidance" but also the fact that "the criminal justice system is seen as increasingly time consuming" (p. 143). This latter explanation may be the strongest one, since there is no long-term decline in notices, which could have occurred had the institution really "moved away" from enforcement altogether. In fact, HSE staff seems to have become increasingly focused and efficient at maximizing their enforcement effect. This is suggested by the percentage of conviction and average penalty per conviction, which are clearly on an upwards trend (Tilindyte 2012, p. 149) – and also by the increase of the ratio of notices per inspections (see Tilindyte 2012 p. 238 – the trend has strengthened since then, with inspections decreasing rapidly and notices remaining at a rather high level). The data provides significant support for the claim of stronger targeting.

Thus, the assertion that there has been a collapse in enforcement, and a trend of "under-enforcement" (Tombs and Whyte 2008, p. 8) does not fully hold up to scrutiny – even less, we would argue, the same authors' even more radical claim of "regulatory surrender" (2010). Still, their point that there has been a tendency to resort less to criminal prosecutions is held up by data – and is not disputed by HSE or any other scholar. In other words, it is "a feature and not a but". The question is whether such an approach – less inspection visits overall, more risk-focused targeting, emphasis on guidance and advice, risk-based enforcement with limited use of prosecution but substantial use of notices – delivers positive results or not for the country at large.

As said, critics of the evolutions in HSE (and LAs) practices tend to use dramatic language – "neutered watchdog", "regulatory surrender", "safety crimes". In a way, this can be understood as no more than the counterparty to "better regulation" slogans decrying "red tape", "stifling burden" and the like. Let us look, then, at the substance of what these critics say. First, they make a number of unproven assertions (or claims resting on very shaky ground at best). Hearings of the Parliament's Committee on Work and Pensions (2008) thus list claims that there is a "correlation between the decline in the inspection rate and increases in fatal injuries", basing it on data from only one year. In the same hearings, some "argued that it was essential that the rate of inspection was increased and that doing so would ultimately decrease enforcement and prosecution costs", but with as little evidence to support it (p. 27). Likewise, in the introduction to a paper by Tombs and Whyte (2008), the editorialists assert that "most safety crimes are either undetected or filtered out from official channels of resolution" – which, in fact, remains unproven through their paper (which does demonstrate other things, but not this claim). The authors then declare that "deaths and injuries suffered at work usually result from infractions of the criminal law" (p. 2), but give no data to back up this statement. They then proceed to claim that their estimates of "deaths and injuries caused by working" are "more accurate", but this claim of "accuracy" is highly debatable (they simply look at another "perimeter"). They also claim that "this process of decriminalisation is reaching crisis point" (*ibid.*) – but show no evidence of it.

In fact, their paper is remarkable for the absence of demonstration of any *trend* in work-related deaths and injuries, or any attempt at comparing their level in Britain with that found in other countries. Their 2010 work (*Regulatory Surrender*), though far longer, similarly avoids the question of trends, beyond claiming that official statistics are "not credible". It makes no effort to look at other sources of statistical information (for, we would contend, they would *not* support their claims). Finally, and this point we will discuss in more details, they pretend to "discover" some data that would have been "hidden" – in fact only reflecting different possible definitions of "work-related" – and similarly claim to "reveal" the unreliability of data (RIDDOR reported number of work-related accidents) that everyone (including HSE management) knows very well to be

incomplete[515]. Tombs and Whyte also suggest to include work-related traffic accidents in the total number of work-related deaths, which indeed increases the total but (absent consideration of chronological trends, of cross-country comparisons) tells us nothing about the performance of the health and safety system (even less so, considering that HSE or LA inspectors, or their counterparts in other countries, are not primarily responsible for road safety, which is under the supervision of other institutions). Finally, they criticize official data for excluding "deaths to members of the public sustained through working environments which are recorded by the HSE" (p. 2). While true, this is again a question of definition. In order to compare across countries OSH data, definitions need to remain comparable. The real question, i.e. whether the situation in Britain has gotten worse, or is worse than in other (more "strictly enforcing") countries, is never addressed.

In fact, the problem of the different definitions of "occupational injuries and deaths" is well known, and has been considered in a number of studies (Feyer *et al.* 2001, Australia's National Occupational Health and Safety Commission 2004, HSE 2014, US Bureau of Labor Statistics 2014 etc.). These differences can sometimes be reconciled so that comparisons are possible, but not always – in any case, they reflect different methodologies rather than some "plot" to hide the scale of a given phenomenon. Some countries (e.g. Germany) include work-related traffic-accidents, but this is more the exception than the rule. Since Britain (and Germany) both are EU members, the European Statistics Office (Eurostat) compiles statistics on key OSH indicators which are harmonized, i.e. where data has been recalculated to conform to a uniform definition. Such data allows easy comparisons between countries. It is noteworthy that none of the strident critics of changes in health and safety regulation in Britain has apparently seemed worthwhile to consider it. These show that the UK has one of the lowest rates of traffic-related deaths[516], strongly questioning the strength of Tombes and Whyte's claim that the number of work-related traffic accidents would be shockingly high. Overall, the UK has a life expectancy that is slightly below the EU average[517] - but not because of violent deaths, where it has consistently among the best EU indicators. Thus, while Tombs and Whyte's are probably right that broadly-defined "work-related deaths" are indeed higher than deaths caused by violent crime, this says nothing about the *evolution* of work-related deaths (however broadly defined), nor about the *relative importance* of this problem compared to other causes of premature deaths in the UK. To be fair, they make a valid point that "the ongoing moral panic that characterises social responses to most 'mainstream' violent crime" (p. 11) makes a strange contrast to the relative indifference to work-related deaths that are, depending on the definition taken, nearly as frequent, or maybe even more frequent (if we take the most expansive definition). They do not, however, prove the importance of work-related deaths in Britain. It may just be that the problem is opposite, i.e. that "mainstream" violent crime is emphasized far too much, and would deserve far less attention, and non-violent causes of death far more.

ii. Comparing health and safety outcomes: Germany and Great Britain

If we want to look beyond *outputs*, and also not base our assessment on assumptions but, as much as possible, on facts, considering Britain's health and safety outcomes *in a comparative perspective* seems unavoidable. It

---

[515] Reporting of Injuries, Diseases and Dangerous Occurences Regulations (RIDDOR) is the official, mandatory system through which employers have to report to regulators any significant incident. Because RIDDOR reports frequently (for injuries) or even systematically (for fatalities) trigger inspections, employers have an incentive to under-report. While this is generally impossible for deaths (which end up criminally investigated in most cases), it is quite feasible for injuries, and results in very significant under-reporting. This is a very well known problem (and has similarities in many countries), and is e.g. covered by Tilindyte (2012), pp. 122-123. In its summary yearly statistics the HSE, well aware of this issue, reports both RIDDOR *and* the more reliable data from the Labour Force Survey (see e.g. http://www.hse.gov.uk/statistics/overall/hssh1314.pdf for the 2013-2014 statistical summary, reporting both figures).

[516] See summary document by Eurostat on causes of death in Europe: http://ec.europa.eu/eurostat/statistics-explained/index.php/Causes_of_death_statistics

[517] See Eurostat summary document on life expectancy: http://ec.europa.eu/eurostat/statistics-explained/index.php/Mortality_and_life_expectancy_statistics

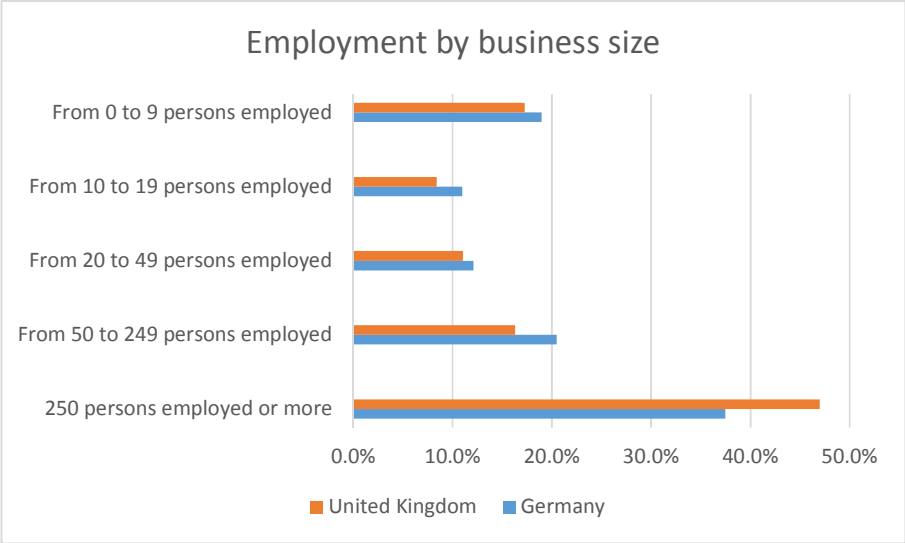
is, however, not so easily done. As Hawkins (2002) pointed out, determining and measuring the effects of health and safety inspections and enforcement is very difficult – for "it is very difficult for regulatory agencies to exhibit their effectiveness in terms of the numbers of injuries or deaths that did not occur" (p. 8). Even if one manages to overcome this problem by looking not at absolute numbers but at trends, remains the near-intractable issue of *attribution*: not only is it difficult to "see whether situations are improving or not", but even then "how much credit do you give to proper regulation, proper firmness, proper inspections?" (pp. 8-9). Overall, "reductions in occupational death, disease, and injury are hard to attribute to changes in law or its enforcement without having a clear sense of the extent to which other influences, such as technological or economic shifts (…) have contributed to any observed effects" (p. 9). These "problems are exacerbated where events are rare, or where latency problems make comparative analysis especially hard" (p. 10), as is the case with many cases of occupational-related illness, which can have decades of latency compared to the work situation which caused them.

We believe that these problems can, to some extent, be alleviated. First, by focusing our comparative research on indicators that have the least latency, and are most easily comparable – in particular fatal occupational accidents. While looking only at one single indicator does not allow to fully reflect on the performance of an institution or a system, we will try and see whether the findings hold constant when considering other (less reliable) indicators. If they do so, we will conclude that there is at least some level of likelihood that the findings do reflect the actual performance of the systems. Second, we will try and filter out the effects of other factors by comparing the outcomes of two countries that are sufficiently similar in terms of economic and social structures, and work methods, and see if there appears to be significant difference in inspection practices, and in outcomes.

The complex issue of "comparability"

The extent to which two jurisdictions can be said to be "comparable" or "similar" is difficult to define precisely, because of the number of variables involved, and the difficulty to assess the relative weight of these variables. Differences between Britain and Germany are quite real – but the average differences are in many cases smaller than the intra-country differences (e.g. on GDP per capita, where the gap between South-East and North England or between Bavaria and Sachsen-Anhalt is larger than the aggregate Britain-Germany difference). If we had the ambition to build a rigorous mathematical model, we would logically also have to try and quantify the most important differences, and attempt to correct for them. Since we have settled, for reasons we exposed above, for a more modest (and, to our mind, meaningful) approach, we will not do so. Germany and Britain are two advanced economies, two of the earliest to have industrialized, two of the earliest to have had a meaningful regulatory system for health and safety (even though Britain was far earlier on the two accounts). The population size and enterprise population is also sufficiently similar (Germany larger on both accounts, but clearly within the same "group" among EU countries). Looking for factors that could plausibly affect health and safety, climate and physical geography are also sufficiently similar (levels of flooding do differ, for instance, but again more between different regions of the country than overall, with the two countries in the same broad climatic zone). If there are meaningful differences that could bias the results, they could mostly be in two areas: employment structure (sectors, size of businesses etc.), and the more difficult to pinpoint field of "culture" and "social norms". While the first question can be relatively easily addressed by looking up the Eurostat website, the second is of course far more complex and disputed.

Let us start by employment structure. According to Eurostat data[518], in 2014 the two countries had active populations around 41.7 million (Germany) and 32.6 million (UK[519]) and employed populations of respectively 39.9 and 30.6 million. The structure of employment by size of business is not strikingly dissimilar, but does show the UK having significantly more employees in large businesses, and Germany more in middle-size ones (the famous "*Mittelstand*").



Unfortunately, neither the British HSE statistical publications nor the German *Sicherheit und Gesundheit bei der Arbeit* (SuGA) reports[520] present occupational accidents differentiated by enterprise size. While this data may be available, it is relatively peripheral to our topic, so we did not investigate further. Given that neither country appears to consider it a meaningful type of disaggregation, and considering that the difference in distribution by size is limited, we can consider its potential effects to be rather negligible. At most, the UK structure might give it a slight advantage, if we assume that large businesses have stronger safety procedures (which is not always true, even though it is often "received wisdom").

The distribution by enterprise sector is somewhat more different, but again is relatively negligible in its effects. As is well known, Germany has one of the highest rates of workers in manufacturing in Europe, while Britain has evolved far more towards a "service economy". Manufacturing, however, is not anymore the sector where work accidents (and in particular fatal ones) are the most frequent[521]. In Britain, fatal injury rates in manufacturing are only slightly above the average for the entire economy, nearly 20 times lower than in agriculture, more than 8 times lower than in waste and recycling, more than 3 times lower than in construction (HSE 2015, p. 3). Germany has a different methodology for calculating fatal accident rates (it includes transportation accidents related to work), and Eurostat does not have the same level of disaggregation, thus it is difficult to fully compare the figures, but the SuGA report for 2013 (BAuA 2014) similarly shows

[518] See Eurostat Labour Force Survey data at: http://ec.europa.eu/eurostat/web/lfs/data/database and Structural Business Statistics tables available at: http://ec.europa.eu/eurostat/web/structural-business-statistics/data/main-tables

[519] Eurostat statistics are for the UK and not for Britain only. The inclusion of Northern Ireland, given its small population and economic weight, is unlikely to change much to the overall picture.

[520] See BAuA 2014 (available at: http://www.baua.de/de/Publikationen/Fachbeitraege/Suga-2013.html;jsessionid=F408E48F5081F0A5F2FE529FC070E1E2.1_cid343) and HSE 2015 (available at: http://www.hse.gov.uk/statistics/pdf/fatalinjuries.pdf)

[521] Assuming it ever really was. As we have discussed in the first Chapter, the high accident rate e.g. in agriculture was for many decades left unaddressed, probably because of different risk perceptions ("traditional" activity against "disruptive" one). Still, manufacturing had very high accident rates a century ago, and even more two centuries ago, as seen in that same Chapter. This is not anymore the case.

manufacturing as slightly above average only, 40% less than in waste and recycling, twice lower than construction, nearly 2.5 times lower than in agriculture[522]. In conclusion, the differences in employed population distribution appear unlikely to result in a large effect on work-related accidents, and particularly fatal one. It may have a *small* effect, though – so the comparison can be valid, but not if it comes to drawing conclusions from small variations. Unfortunately, while Eurostat does collect and compute accident rates for each economic sector separately, the full disaggregation is not available online (only broad categories are available, lumping together all agricultural and industrial sectors). The differences between countries in terms of economic structure are, in any case, accounted and corrected for in the standardized injury/fatality incidence rates compiled by Eurostat, which are therefore fully comparable and can be taken to reflect the frequency of such injuries *irrespective* of differences in distribution of employed populations between sectors[523].

What, then, to think of the "cultural difference" between the two? Many sweeping generalities and unfounded assumptions circulate on the different cultures and their characteristics (and effects), but few (if any) are seen to hold when confronted with careful examination of facts and data. As Chang (2007) reminds us, the stereotypes that cast Germans are more hard working, more respectful of rules, were reversed in the 19th century, when British writers found them to be "lazy" and "lying" (pp. 179-183). The recent and still unfolding Volkswagen emissions scandal is, from this perspective, a welcome reminder that stereotypes have little to do with reality. From our perspective of making "modest" comparisons, not attempting to model or draw strict quantitative inferences, the similarities between Britain and Germany appear large enough. To many considering them from afar, certainly, they are significant: two societies in North-Western Europe, wealth, with a long history of state building, public administration, legal compliance, social services etc. Anecdotal evidence underlines the proximity, e.g. data on fatal traffic accidents. Out of many indicators, we selected this one because it reflects to a significant extent on attitudes towards compliance with rule and safety issues. Because the two countries have (once again) sufficiently similar size, population density, quality of roads, wealth etc., other factors that could lead to large discrepancies in traffic fatalities can be discounted for a first approximation. Both Britain and Germany end up belonging to the same group of low-fatality countries, with Britain performing slightly better, whichever indicator is used (fatalities per population, per motor vehicles, or per passenger-kilometres)[524]. Thus, again, we can for now hold that the similarities between the two countries are sufficient to allow for meaningful comparison, though certainly not for drawing excessive conclusions from small variations.

Health and Safety outcomes in Britain and Germany

Lets us first look at outcomes in both countries, and to this aim let us focus on the indicator that is the least susceptible to under-reporting (and thus to bias in reporting that could differ from one country to the other depending on specifics of regulations and practices): fatal accidents at work. Non-fatal accidents tend to be under-reported in Britain (as in many other countries), because they may lead to additional inspections, potential liability issues etc., and employers try and incite or coerce their employees in not reporting them

[522] It is likely that the fact that the rankings are similar, but the magnitude of the gaps far smaller in Germany, is linked to the inclusion of work-related traffic accidents. These can be expected to be somewhat more rare e.g. in construction, and more frequent in manufacturing or services, thus reducing the difference between lowest and highest rates. To fully explain the much lower gap between agriculture and the average, one would have to dig deeper in the data, but caution is needed because the total numbers are low anyway (and the agricultural working population quite small in both countries), thus meaning that variations of a very small number of actual cases can yield considerable changes in percentage points.

[523] See full methodological note on European statistics on accidents at work on Eurostat website here: http://ec.europa.eu/eurostat/cache/metadata/en/hsw_acc_work_esms.htm.

[524] See the Wikipedia page on *List of countries by traffic-related death rate*, with data for all the three indicators, available at: https://en.wikipedia.org/wiki/List_of_countries_by_traffic-related_death_rate.

either. By contrast, in Germany, the no-fault insurance system and the multiple control points (e.g. reporting obligations of medical doctors) mean that the reported level is expected to be far closer to the real one[525]. Accidental deaths also avoid the time-lag problem that makes it difficult to link occupational-related illnesses with *current* regulatory practices (as the effects today may often be related to practices one or several decades ago). Thus, even though the Labour Force Survey (LFS) provides more reliable accident figures for Britain (and they are indeed the ones used by both the HSE and Eurostat), fatal accidents appear the most convincingly comparable (also because, for non-fatal accidents, the question of the *severity* of the accidents would have to be taken into account).

As briefly noted above, fatal accident *rates* being very low, a change of a few units from year to year can lead to important changes in percentage points – thus it is important to compare not only rates for one given year, but averages over a longer period. A further difficulty arises when considering *which definition* of the fatal work related incidents rate to take. Eurostat's definition[526] of an accident is "a discrete occurrence in the course of work which leads to physical or mental harm" – and a *fatal* accident is one "which leads to the death of a victim within one year of the accident[527]". Crucially, the definition *since 2008* includes "all accidents in the course of work, whether they happen inside or outside the premises of the employer (…) in public places or during transport (including road traffic accidents or accidents in any other mean of transportation) and at home (such as during teleworking)" (though it excludes "accidents on the way to or from work"). The HSE's definition, by contrast, *excludes* accidents taking place during transport, which results in a markedly different picture (it corresponds to Eurostat practice up to and including 2007).

As pointed out by Tombs and Whyte (2008), the exclusion of traffic accidents involving "at work" vehicles is important – it does not only change the overall magnitude of the problem (which could be without consequence for comparisons, if the change was constant across countries), but in fact also changes the *difference* between British and German fatal injury rates, as well as part of their *evolution*. Because Eurostat still publishes data excluding traffic accidents, we are able to consider both alongside each other, as well as long-term averages. The tables below presents the evolution and averages of these standardized rates, as obtained from the Eurostat website (both excluding *and* including "at work" traffic accidents) – the former corresponding also to the data in the HSE publication *European Comparisons –Summary of UK Performance* (2015)[528] and *Statistics on fatal injuries in the workplace in Great Britain 2015*[529]. As we can see, while Britain performs *overall* better than Germany, the difference is far sharper when excluding traffic accidents, and even ends up slightly reversed in some recent years when including them. Because of changes in Eurostat procedures, there are pre- and post-2008 figures with different references and definitions[530].

Eurostat data *Standardized incidence rates, fatal occupational injuries – excluding traffic- and transport-related*	2008	2009	2010	2011	2012	2013	**1998-2007**	**2008-2013**	*1998-2013* [531]
Great Britain	0.59	0.59	0.69	0.73	0.58	0.51	**1.4**	**0.62**	*1.11*
Germany	1.11	0.66	0.81	0.94	0.9	0.81	**2.1**	**0.87**	*1.66*

[525] See Tilindyte 2012 pp. 122-123 (Britain) and 181-182 (Germany). On fatal accidents see pp. 121-122 for Britain specifically.
[526] See: http://ec.europa.eu/eurostat/cache/metadata/en/hsw_acc_work_esms.htm.
[527] Considerable harmonization work is needed because of this, as different Member States have very different durations being considered for their own definition of "fatal accident".
[528] Available on the HSE website, statistics section, at: http://www.hse.gov.uk/statistics/european/european-comparisons.pdf
[529] Available on the HSE website, statistics section, at: http://www.hse.gov.uk/statistics/pdf/fatalinjuries.pdf
[530] All standardised incidence rates for fatal accidents are per 100,000 workers. For the methodology, see: *European Statistics on Accidents at Work (ESAW) – Summary Methodology,* 2013, Eurostat – available at: http://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/KS-RA-12-102. Pre-2008 data excludes traffic and transport accidents (even if work-related or during work time), while data since 2008 includes them (but Eurostat also offers tables for incidence rate *without* these).
[531] Due to change in methodology in 2008, this average is only for informational purposes. Last line is for EU 27.

EU 15	1.83	1.64	1.58	1.43	1.3	1.19	**2.4**	**1.5**	*2.04*
EU 28 *(EU 27 until 2008 included)*	2.31	1.94	1.87	1.59	1.46	1.3	**2.6**	**1.63**	*2.26*
Eurostat data *Standardized incidence rates, fatal occupational injuries – including traffic- and transport-related*	2008	2009	2010	2011	2012	2013		**2008-2013**	
Great Britain	1.02	1.55	1.61	1.8	1.52	2.05		**1.59**	
Germany	2.67	1.4	1.58	1.59	1.54	1.29		**1.68**	
EU 15	2.68	2.31	2.35	2.42	2.33	2.17		**2.38**	
EU 28 *(EU 27 for 2008)*	3.1	2.52	2.61	2.65	2.44	2.22		**2.59**	

In any event, Great Britain's performance in occupational safety and health, using fatal injuries as a proxy, is way better than the EU average, even using the EU-15 group only, i.e. the "oldest" (and wealthiest) members. It is vastly better, for instance, than France (2008-2012 average of 3.83 including traffic accidents or 2.62 excluding them) or Italy (3.64 and 1.52 respectively over the same period), a point worth coming back to later, considering the sharp differences in OSH inspections between Britain and these two countries (France in particular). The comparison with Germany is, however, slightly more complex. If we consider the definition *excluding* traffic accidents (which is the one used in Britain by the HSE and corresponds to its mandate, and is also the one for which longer-term data is available), then the gap is constantly in favour of Britain, and it has remained remarkably constant (33% lower over 1998-2013, 30% lower for 2008-2013, 37% in 2013).

In recent years, however, if we still consider this same definition, Britain's performance appears to have plateaued (and even worsens in 2013, with a Eurostat footnote indicating a "change in data series", meaning the trend should be checked again in 2014), whereas Germany's improved markedly between the beginning and the end of the period considered. While their performance is essentially similar on average over 2008-2013 (Britain's rate being 5% lower overall, a slight edge only), there are important swings from one year to the next, and Britain's rate goes from 60% lower to 60% higher than Germany's. Overall, swings in data including traffic accidents seems to be substantially stronger, possibly linked to the far higher number of factors that could influence the overall rate, and the potential for "catastrophic" road accidents having an influence on the data.

For a variety of reasons, we have concluded that the more meaningful figure to compare the effectiveness of the inspection and enforcement system in achieving good OSH outcomes is the incidence rate of fatal accidents *excluding* traffic-related accidents. First, this indicator is available on a longer timeframe, which is important because of the high year-on-year variability of the rate (given that fatal accidents are anyway rare). Second, and crucially, it corresponds to the sphere of responsibilities of the HSE and LA inspections in Britain (see below). Third, *even* if we consider the least-favourable indicator (*including* road traffic accidents), Britain's performance remains at least as good as Germany's, while relying on a far smaller number of inspections. If we consider data *excluding* traffic accidents, not only has Britain long had among the best OSH performance in Europe, but its edge over Germany has held in spite of strong improvements in Germany's performance. – but also suggest that this edge has been eroding, not because Britain's performance worsened, but because Germany's improved. Over the period 2008-2013, Britain retains on average the lowest fatal injury rate in the EU, with the Netherlands, Slovakia and Germany coming close (in that order). Finland, Denmark and Sweden also rank among very good performers, but with somewhat worse data. In the period 1998-2008, the best performance was Sweden's, followed by Britain, the Netherlands and Finland. Thus, Britain has confirmed its excellent performance over the long term, being one of the very best in Europe, but Germany has improved

its relative ranking – in a period where, in fact, inspections decreased in numbers, as we will see in the next section.

If we use the incidence rate *including* traffic-related accidents, the best performer is Greece (which is likely to reflect the collapse in economic activity since 2008, particularly in high-risk sectors for OSH such as construction), followed by the Netherlands, the United Kingdom (data for the whole of the UK rather than only Britain), Finland and Germany. Thus, changing the indicator would not meaningfully affect our conclusions.[532].

The significantly stronger difference between British and German performance when excluding traffic accidents, compared to when they are included, may reflect structural differences, with a stronger share of activities involving more intensive use of transportation during work. It also most likely reflects far less emphasis on reducing this kind of occupational accidents, which both justifies and questions HSE's reporting data excluding them. On the one hand, as HSE is *not* responsible for investigating such fatalities as per current legislation, it can be understood that they are outside of its remit, and that prevention of such fatalities is the province of other state authorities (those responsible for traffic safety more broadly). On the other hand, anecdotal experience suggests that there can be a lot of employer pressure on employees driving to "cut corners" in order to meet schedules and targets, and these should be recognized as "occupational", and addressed as such (by HSE and LA EHOs, as relevant). Overall, while this gap does not really affect Britain's excellent performance in OSH, and only marginally affects Britain's ranking, the continued exclusion of traffic- and transportation-related occupational injuries and deaths may be misleading, and make it more difficult to work effectively on reducing them. At the same time, this very difference actually goes a great way to demonstrate the HSE's (and LA inspectors') effectiveness: whereas Britain has established a major gap in its favour when it comes to occupational safety "on premises" (where their activities are focused), this edge is far less pronounced when traffic-related accidents are included, which suggests that, on traffic-related occupational accidents *alone,* Britain's performance is clearly worse than Germany's. While this suggests that it would be important to enlarge the scope of OSH supervision to further reduce accidents in Britain, it also demonstrates very clearly that the way HSE and LAs work is highly effective, and that the far lower number of inspections they conduct does not negatively affect this performance.

Health and Safety inspection practices in Britain and Germany

We can thus conclude from the above that, at least on the metric considered most reliable and most easily comparable (fatal occupational accidents), Britain has been performing generally significantly better than Germany, though the latter has improved its results in recent years. British performance is worse in areas where OSH inspectors are not involved (traffic-related accidents), but this only strengthens the evidence that these inspectors (and their institutions) are highly efficient and effective in their domain. Because of considerable problems with their reliability, we will not try and complement these data points with a comparison of non-fatal injuries (a quick look at Eurostat data shows massive under-reporting in a number of countries, making the whole data set unfit for use[533]). A glimpse at occupational health statistics would be useful, to balance the "short term" perspective (injuries) with the "long term" one (diseases), but again Labour

[532] Over 1998-2007, Greece's performance is far worse, suggesting that its excellent rating post-2008 may be (in a paradox frequently observed in OSH) linked to the brutal economic crisis, leading to a massive slump in some high-risk areas (e.g. construction).

[533] A look at the Eurostat tables for injuries resulting in more than 3 work days lost shows that the average rate (2008-2012) was lowest in Romania, Bulgaria, Latvia and Lithuania, with Romania's rate less than 1/30th of the EU average, and Germany's around 15% above EU average (Germany's system, as noted above, makes it likelier that work-related injuries will be correctly recorded). By contrast, on fatal injuries, Romania has the EU's worst performance (more than twice the average incidence), followed by Lithuania. Clearly, under-reporting is massive, even when Labour Force Surveys are used (work-related injuries are just not "perceived", and/or are actively hidden, and/or workers are simply unable to take days off for fear of losing their job, etc.).

Force Survey (LFS) data presents glaring inconsistencies, with some good performance pointing clearly to under-reporting, but certainly not *all* good performance. Let us just suggest here that, considering Britain's excellent record on fatal injuries, there is at least some plausibility that its good record on work-related health problems as reported by the LFS may reflect reality rather than under-reporting[534].

The second part of our comparison will now focus on OSH inspection practices in both countries – their numbers, targeting methods, and "style". As Tilindyte (2012) has shown, the structures and practices in Britain and Germany differ significantly, and it is precisely these differences, as well as the considerable distance in terms of inspections frequency, that matter to us. The structural differences mean, however, that comparisons are not as straightforward as one may wish. OSH inspections in Germany are conducted by two sets of bodies: state officials working for the federated states (*Länder*) in the Enterprise Supervision services (*Gewerbeaufsicht*), and employees of the mandatory insurance providers (*gesetzliche Unfallversicherungen*). While the latter focus exclusively on occupational safety and health issues (checking compliance with technical norms that are meant to reduce accidents and illness, and thus insurance payments), the former have a mandate that is less precisely defined. In some *Länder,* the *Gewerbeaufsicht* are assigned a number of non-OSH related missions, in some their organization has been devolved to the local (municipal) level (Tilindyte 2012, pp. 166-167). In some cases, the remit of these inspections includes consumer or environmental protection (*ibid.*, p. 175). These inspectors also control provisions of legislation relating to child labour, work time etc. – but *not* "provisions of collective agreements, and they do not enforce legislation in relation to social security and employment contracts, such as the payment of wages or dismissal"[535]. Thus, with the exception of cases where inspectors are responsible for consumer or environmental law, or market surveillance[536], their remit is roughly comparable to the HSE's (somewhat broader in terms of labour legislation, somewhat narrower in terms of health and safety, where the HSE is mandated to look in a holistic way and not only at the occupational perspective). In some *Länder*, the supervision of state laws on occupational safety and health has also been entrusted to the mandatory insurers, at least in some sectors[537] – but this has no incidence on the total number of inspections, only on who conducts them.

Thus, for Germany, in order to assess the total number of inspections we have to consider all the visits conducted by the mandatory insurers, as well as *most* (but possibly not all) of the visits conducted by the *Gewerbeaufsicht.* In fact, it is unlikely that some visits are *exclusively* focused on non-OSH issues, and most likely that all visits incorporate at least some elements of OSH, thus there is sufficient ground to assume that all of these visits can be counted as OSH-related. In any case, as we will see, even if we applied some discount, the gap between Britain and Germany would remain considerable.

The data from Britain is somewhat simpler to interpret, but also carries some degree of uncertainty at the local level, this time not because what is counted as OSH inspections may incorporate non-OSH visits (though this may also be the case), but because it is possible that environmental health officers (EHOs) conducting non-OSH visits (e.g. focusing on hygiene) also look "on the side" at OSH issues, thus conducting a form of monitoring that may improve the overall coverage and ability to identify risks. This is, in fact, "a feature, not a but" of the system – EHOs have a broad set of skills and competences, and a broad mandate, allowing them to cover inter-related risks in several areas during one visit, and joint inspection visits are encouraged, as we have seen above. In other words, the official count of OSH visits by LAs may be to some extent an

[534] See HSE 2014, *European Comparisons* – p. 4.
[535] ILO's summary page on *Labour Inspection Structure and Organization* in Germany – available at: http://www.ilo.org/labadmin/info/WCMS_209470/lang--en/index.htm
[536] See ILO, *ibid.*
[537] See ILO, *ibid.*

underestimation – and it is difficult to put a figure on how significant it may be. Again, we will see that applying some correction to these numbers would not change the overall conclusion.

Germany (SuGA 2013, pp.286, 292-293)		Number of enterprises visited	Number of inspection visits
Länder inspections	2011	123252	297917
	2012	110207	267008
	2013	99999	242503
Mandatory insurers	2011	353360	639641
	2012	337345	603483
	2013	297941	599605
Total	2011	476612	937558
	2012	447552	870491
	2013	397940	842108

Great-Britain	Local authorities	HSE	**Total**
2006-07	*200000*	50000	*250000*
2009-10	196000	45000	241000
2010-11	194200		
2011-12	151000	*37000*	*188000*
2012-13	106200	33000	139200
2013-14	86900	30000	116900

Sources: HELA Paper H15/01, LAE1 return forms, HSE annual reports, Löfstedt 2011, "Focus on Enforcement" data for 2012-13. Since HSE does not publish aggregate inspections data, the HSE column represents the author's own extrapolations based on incomplete data, experts opinions (direct interviews with OSH enforcement officials in Britain) and breakdown of inspections in earlier years for which full data is available. Numbers in italics denote author's extrapolations based on trends (no data source available for that year).

The above tables show that there are currently more than 7 times more OSH inspections in Germany than in Britain, and that even in earlier years (when OSH inspections were more frequent in Britain), the ratio was around 3.5 to 4 times more in Germany. This is quite a striking difference. Of course, the number of active businesses (and of establishments to be visited), and the active population, are also different, and significantly larger in Germany – hence it is important to look at inspection *rates* rather than absolute numbers. Ideally, like Eurostat data, the inspection rates should be normalized correcting for different economic structures, but the British data is not sufficiently differentiated to allow for this (SuGA reports give data by economic sector and would in principle allow for such corrections). We will have to accept this limitation, considering (as indicated above) that the two countries' economic structures are, though by no means identical, sufficiently close.

Adjusting for population requires to take a decision as to which population to use (workers or establishments), and (if the latter) which definition of establishment, enterprise (or business) population to adopt. This is all the more important as Britain has seen an important shift from salaried employment to self-employment, i.e. an important rise in the number of individual businesses, most of which are *not* being controlled by the HSE or by LAs[538]. There is no perfect choice in this matter, so we will present below the different possibilities, as well as the results obtained with each of them. The best variable for adjustment would be the number of "establishments (premises) under supervision", but this is not available in either country (though some reports in the UK present estimates, which we also refer to below).

	Great Britain	Germany	ratio Germany/Britain
Total population (millions)	62.70	82.70	1.32
Active population (millions)	33.01	44.70	1.35
Employed population (millions	31.09	42.70	1.37
Persons employed in private businesses	25,354,000	31,914,340	1.26
Total number of businesses	5,272,530	3,629,666	0.69
Number of businesses with 10 employees and more	236,830	339,087	1.43
Persons employed in non-zero employees businesses	20,996,000	29,914,278	1.42
Persons employed in businesses with 10 employees or more	17,093,000	24,176,805	1.41
average ratio (excluding total number of businesses)			1.37

(Sources:

Great Britain: Office for National Statistics (ONS): Labour Market Statistics – September 2015 release; ONS: Population Estimates for the UK, England and Wales, Scotland and Northern Ireland, mid-2014 release; Department for Business, Innovation and Skills (BIS): Business Population Estimates for the UK and regions, 2015

[538] Löfstedt (2011) has recommended *not* inspecting businesses with no employees except when their type of activity could pose a significant risk to outsiders. This policy, while not having the force of law, is widely implemented (and in fact was most probably widely implemented already before the report).

Germany: *Statistiches Bundesamt* website: *Konjonkturindikatoren, Volkswirtschaftliche Gesamtrechnungen* and *Arbeitsmarkt*; *Statistiches Bundesamt* online database "Genesis": tables 52111-0001 and 52111-0014[539])

If we exclude the total number of businesses, which is much higher in Britain due to the far larger number of self-employed (as mentioned above), the ratio between Germany and Britain is relatively stable, mostly between 1.3 and 1.4. Since most of these self-employed fall outside of the HSE's and LA's OSH inspections remit, we decided not to count this number and consider only the other ratios. The rationale for considering specifically the number of workers in businesses with at least one employee and with more than 10 employees is that they tend to be those (particularly those with at least one employee, as opposed to purely individual entrepreneurs) on which OSH inspections focus. Overall, excluding the total number of businesses, the smallest gap is on the total number of persons employed in private businesses, the largest is in the number of businesses with 10 employees or more. As indicated above, there is another number which would be most relevant to consider, the number of *premises under supervision* (i.e. at least potentially covered by inspections). Unfortunately, this number only exists (as estimate) for Britain. As indicated in the BRE report *Improving Outcomes from Health and Safety* (2008), there were over 2.5 million such premises in 2008 (BRE estimate), with over 26 million workers (HSE estimate) (BRE 2008, p. 59). In Germany, the SuGA 2013 (p. 291) report (BAuA 2014) indicates over 3.2 million businesses and 31.5 million workers as reported by the mandatory insurers, which *may* correspond to the same definition (approximately at least, though the number of *premises* is likely to be higher than that of *businesses*) – but the *Gewerbeaufsicht* may have another "universe" of premises under supervision. Using these numbers, the respective ratios would be 1.28 and 1.21. This would suggest a *smaller* gap between the scope of supervision in Britain and Germany, and actually *strengthen* the case that OSH inspections in Britain are considerably less frequent. We thus decided to keep a ratio of 1.37 for further calculations, which corresponds to that of the employed population.

	number of OSH inspections			
	Germany	Great Britain	ratio	ratio adjusted for population
2006	*1100000*	*250000*	4.4	3.2
2009	*980000*	241000	4.1	3.0
2011	937558	*188000*	5.0	3.6
2013	842108	116900	7.2	5.3

(N.B.: data in italics are estimates – for Great Britain, author's own extrapolations based on available data – for Germany, based on Tilindyte 2012 p. 177)

Once adjusted for the difference in population, we see that for a number of years the frequency of OSH inspections adjusted for population was 3 or more times higher in Germany. In the most recent years, the gap has considerably increased, and the inspections rate is now above 5 times higher in Germany. While inspections have overall become less frequent in both countries, the gap has remained, and recently (in line

[539] Reconciliation of data for Germany was very difficult, with various sections and tables of the Federal Statistics Office giving different totals with apparently similar definitions. These two tables were taken as being the most reliable, and also match best the data available p. 291 in the SuGA report (BAuA 2014), which is based on mandatory insurers reports. It should be noted that, for Britain, ONS and BIS have different statistics, apparently based on different definitions and sources. We have used the BIS reports as being the most up to date and corresponding best to our definitions. *Pro memoria,* in Germany there are also statistics on the number of *Betriebe* distinct from those on *Unternehemen,* i.e. a broader "establishments" definition including public sector ones, with a total of 3,835,716 – we did not use it because it differs from the UK definitions, and would change little to the ratio. There also are discrepancies between "employed" and "subject to social security contributions" numbers in some cases. The numbers quoted in the table are "subject to social security contributions" as per the Federal Statistics Office tables referenced.

with post-2010 reforms in Britain) increased. Even if we discount this recent increase (since, as we will discuss below, the reforms may in some respects have gone further than what risk-based principles would recommend), the long-term ratio of more than 3 times is very high. As we indicated several times above, considering the many ways in which the two countries differ, a small variation in inspections frequency (or in outcomes) could be discounted. Three times more frequent inspections (and more), for outcomes that have been in the long term generally far worse and (at best, if taking the most favourable indicator) are now equivalent appears like a strong indication that inspections and enforcement practices in Germany are clearly less efficient than in Britain[540]. This example clearly shows that the naïve assumption that "more inspections" will automatically mean "better outcomes" is mistaken, it also leads to ask *what* makes British regulatory delivery so much more effective.

iii. Explaining effectiveness and efficiency – risk-based approaches vindicated?

Conceivably, at least some part of this difference in efficiency could be explained by British inspectors having access to more effective deterrence tools than German ones – or having a more adequate "enforcement pyramid", making escalation more credible and therefore more effective at increasing compliance. Tilindyte (2012) suggests that this may be the case, after an in-depth review of the enforcement options available to inspectors in both countries, considering not only the letter of the law but the actual practice, and the extent to which theoretically available sanctions are used in practice (pp. 230-274). On administrative sanctions, her conclusion, on balance, is that administrative sanctions are relatively more frequently used in Britain than in Germany, and that although the HSE only has notices (and, to some extent, license terms modification) at its disposal, the publicity of the notices (public registers) adds a powerful "naming and shaming" effect. By contrast, German inspectors have in theory more varied and powerful administrative sanctioning tools, but in practice (for reasons both of complexity of procedures, of ignorance of the option to impose *corporate* rather than *individual* sanctions, and of "culture") use them very rarely (pp. 270-271). Concerning criminal sanctions, she likewise concludes that both probability and potential severity are somewhat higher in Britain – and that the lack of corporate liability could be a significant limitation in Germany (pp. 269-270). These relative weaknesses may indeed *to some extent* contribute to making German inspections less effective in terms of dissuasion – in particular, existing research indeed suggests that negative publicity may be a stronger driver than sanctions themselves, and the legal strength of the British "notices" is quite significant. The lack of corporate liability in Germany is clearly a significant problem. Nonetheless, these are unlikely to be sufficient to explain how an inspection rate more than three times lower results in equivalent or better outcomes in Britain, particularly considering that deterrence is (as we have seen in the second Chapter) a relatively weak compliance driver.

The answers can be found by looking more closely at data and practices, in particular concerning *targeting,* the way advice and guidance are *structured*, and the overall organization of the system. On the first point, Tilindyte presents (*ibid.,* pp. 238-239) historical data on inspections, "deficiencies" identified (for Germany), and enforcement measures. Though the data is partial (only HSE FOD for Britain, and only *Gewerbeaufsicht* for Germany), it is enlightening. While, for the periods considered, HSE FOD conducts only 12-13% as many inspections as the *Gewerbeaufsicht*, it issues 55-66% as many notices, and initiates from 2 to 7 times as many prosecutions (but the German authorities also have the ability to impose direct administrative penalties). In other words, German inspectorates visit far more premises, but seem to find problems that are sufficient to warrant enforcement action only very rarely. This should be considered also in light of the number of "OSH

[540] One could even use this as a possible indication that OSH outcomes have no relation at all to inspections and enforcement practices. For the many reasons exposed in earlier chapters and sections, we believe this would be an excessive claim – but surely this example shows once again that there is no direct correlation between more inspections and better outcomes.

deficiencies" identified in Germany, on average 1.5-1.6 per inspection. Thus, inspectors find a vast number of "deficiencies", but consider most of these to be too mundane for enforcement, or enforcement not to be the adequate response. There are several ways to interpret such findings, including that this may represent "reluctance" to adopt a formal approach (Tilindyte 2012, p. 239) – but there is also a distinct possibility that this reflects a "net cast far too wide". German inspectors may be over-inspecting, i.e. visiting a large number of premises where the risk level is low, then identify "deficiencies[541]" that reflect a large set of very prescriptive and detailed rules (which make non-compliance likely by their very nature), but decline to enforce because they clearly see that it would be disproportionate (and would likely trigger backlash from their own management or from the public, even if their own professional judgement did not discourage them from enforcing).

Interestingly, professional attitudes between German and British inspectors seem to have some similarities, at least in terms of emphasis on advice and "informal" rather than formal enforcement (*ibid.,* pp. 193-194 and 239-240). It does not appear to be a "lack of advice" overall that could explain the relative under-performance of German inspections, or constitute the key difference between the two systems. In fact, the fact that *both* Britain and Germany rank among the very best performers in terms of OSH in the EU (and, more broadly, in the world) suggests that "informal enforcement", rare use of sanctions and emphasis on advice and guidance may be working very well at promoting compliance and safety. Tilindyte quotes the EU's Senior Labour Inspectors Committee (SLIC) as reporting in 2005 as observing a "widespread, seemingly institutionalised, assumption that advice is more effective and preferable to [formal] enforcement" (*ibid.*, p. 239) – and suggests that this, at least, in "tension" with the SLIC principles which, while they foresee the use of "informal enforcement", also put some emphasis on adequate powers and formal enforcement (*ibid.,* pp. 98-99). However, the performance of systems that put heavy emphasis on formal enforcement, such as France's, appears considerably worse. As we have seen, France has a rate of fatal occupational accidents that is very high (2008-2012 average of 3.85, 50% above the EU 15's and 2.5 times more than Britain's or Germany's), with far more inspections than Britain[542] (though not necessarily than Germany's), and an enforcement approach that is well known for being very formal, enforcement-prone and "adversarial"[543]. Clearly, there is a "chicken-and-egg" question as to whether informal, cooperative enforcement fosters effective cooperation, or whether a cooperative climate is what makes informal enforcement possible – but these examples clearly show that equating "intensive, formal enforcement" with "higher effectiveness and compliance" is simply impossible[544].

---

[541] Which do not necessarily amount to an offense – cf. Tilindyte 2012 p. 191.

[542] In 2010, France's Labour Inspection conducted 368,236 "interventions" (inspections, investigations etc.). In addition, each regional Medical Insurance *Caisse* has a corps of controllers (not unlike the dual system in Germany). While no consolidated statistics exist on their numbers or activities, anecdotal evidence suggests their visits are (at least in some sectors) not less frequent than Labour Inspectors'. While in 2011 there were over 2,200 inspectors (broadly defined, as there are two different statutory grades with inspection functions) in the Labour Inspection, an estimate based on data from a few regions suggests there should be close to 1,000 Medical Insurance inspectors at least. Assuming a similar inspection schedule, this would yield between 500 and 600,000 inspections per year, vastly more than in Britain, even though the two countries have a similar working population and number of enterprises. (Source: *L'Inspection du Travail en France en 2011,* officla report to the ILO, available at: http://travail-emploi.gouv.fr/IMG/pdf/Rapport_IT_2011_sans_table.pdf - cf. pp. 63 and 124). Note of course that France's labour inspectors check not only OSH issues but also the full employment legislation, and that this is a major part of their tasks (which may distract their attention from OSH). Still, health and safety are part of every inspection they conduct (in principle at least), and the Medical Insurance inspectors check *only* health and safety.

[543] There are reams of anecdotal evidence but evidence is provided by the same report to the ILO – remarkably little on outcomes, but strong emphasis on inspections and enforcement (pp. 70-102), far more than on advice and guidance (pp. 108-113). Also the same report (pp. 135-137) emphasizes the importance of "protection" (judicial) granted to inspectors in cases of conflicts (including violent ones). Such items are of course missing from e.g. HSE reports, and such situations generally unknown in Britain (or in Germany). The degree of conflict linked to labour inspections in France is particularly high, with inspectors widely known to be strongly politicized and both sides (inspectors and businesses) seeing the other more as "enemies" than "partners".

[544] As indicated, France has more inspections and far worse outcomes than Britain. As sketched out in the first Chapter, the roots of France's labour inspection challenges and practices go back to the 19th century, and combine political, legal, administrative and social

A key difference between the HSE's practices and advice provided by German inspectors appears to be the degree to which, in the UK, advice is provided pre-emptively and with a broad outreach effort, in a way that is designed to be easy to understand and implement. In other words, whereas German inspectors seem to primarily provide *ad hoc* advice based on their findings and their own experience and understanding, the HSE (and, increasingly, LAs) provide *guidance.* This guidance is not provided only when an inspection reveals problems, but proactively, to everyone who requests or looks for it (and HSE and LAs make efforts to ensure that as many people and businesses as possible are aware of this guidance).

As Tilindyte (2012) points out, advising businesses on how to comply with OSH regulations is a *duty* of the German state authorities (p. 190). As she underlines, inspectors "see themselves primarily as consultants, service providers" assisting to achieve compliance, and this is "particularly true for inspectors of the accident insurers" (*ibid.*). In fact, as we have seen, the SLIC's assessment considered that, if anything, the Germany system was too far on the side of "advice" and not enough on the "enforcement" side. Tilindyte further notes "numerous programmes of the individual authorities" that "reflect their aspiration to improve and strengthen education and advice" (p. 191). The "quantity" of advice thus does not seem (at least in recent years) to be an area where there is a meaningful difference between Britain and Germany. To the extent that Tilindyte underlines that evolutions started around 10 years ago led to a strengthening of the "service-oriented", advice component of German OSH structures' activities, it is conceivable that this change may have contributed to the significant improvement in outcomes that we have observed above (with Germany catching up gradually with Britain). Still, it appears that the primary vehicle of advice in German OSH practice remains the inspection visit itself. A (non-exhaustive) look at different government websites covering OSH suggests that significant efforts have been done to improve availability, accessibility and ease-of-use of information, but that much remains to improve – in terms of contents, as in terms of ease-of-use.

The German system's fragmentation may be part of the causes of the problem, meaning that there is a mix of federal and state-level websites to consider, each with a different structure and focus. For this research, we reviewed the information available on two state-level websites (Niedersachsen[545] and Nordrhein-Westfalen[546], the first as an example of a mid-size state with significant manufacturing industry and a

---

elements. Still, it is remarkable that the lack of effectiveness of existing methods has been so little challenged. Even if existing practices of French labour inspectors may have come "in reaction" to "resistance" by businesses, they clearly seem to have the result to *strengthen* them rather than lead to transformations, thus the "chicken-and-egg" question is not useful when it comes to determining a course of action.

[545] See the portal of the *Gewerbeaufischt* for Niedersachsen (Lower Saxony), with the start page available at: http://www.gewerbeaufsicht.niedersachsen.de/. Detailed perusal of the portal shows that there are a number of specific pages e.g. on OSH organization in the workplace, protection against specific hazards etc. The portal also links to a certain number of practical tools – these include in particular a safety inspection check-list focusing on hazardous installations (available at: http://www.gewerbeaufsicht.niedersachsen.de/download/30099/Ausfuellbares_Pruefschema_fuer_Sicherheitsberichte.pdf), and (for the same type of safety visits) safety inspection guidelines (available at: http://www.gewerbeaufsicht.niedersachsen.de/download/81617/Niedersaechsischer_Inspektionsleitfaden_2012_zur_Durchfuehrung_der_jaehrlichen_Vor-Ort-Inspektion_entsprechend_16_Stoerfall-Verordnung.pdf). Practical, easy-to-use guidance is, however, rare. One example is a (rather short and not very user-friendly) flyer on asbestos (http://www.gewerbeaufsicht.niedersachsen.de/download/57589/Flyer_Entsorgung_von_Asbest_Stand_08_2013.pdf), but there is very little guidance overall for construction works, one of the highest risk sectors.

[546] See the portal for occupational safety in Nordrhein Westfalen (North-Rhine-Westfalia), at: http://www.arbeitsschutz.nrw.de/. The portal is well structured and it is easy to find pages helping employers with their risk assessment and explaining their responsibilities (e.g. http://www.arbeitsschutz.nrw.de/themenfelder/arbeitsschutzsystem_gefaehrdungsbeurteilung/index.php and http://www.arbeitsschutz.nrw.de/themenfelder/arbeitsschutzsystem_gefaehrdungsbeurteilung/verantwortung_des_arbeitgebers/index.php). It has practical tips e.g. on lifting heavy weights (see at: http://www.arbeitsschutz.nrw.de/themenfelder/arbeitsplaetze_arbeitsstaetten/heben_und_tragen/index.php) including a brochure (http://www.arbeitsschutz.nrw.de/pdf/themenfelder/sieben_schritte_zum_erfolg.pdf). The latter is, however, rather "conceptual" and targeting managers more than directly "visual" and practical. It has tips for builders (http://www.arbeitsschutz.nrw.de/themenfelder/baustellen/pflichten_des_bauherren/index.php) and a brochure on "safe building" (http://www.arbeitsschutz.nrw.de/pdf/themenfelder/baustellen/Fb_Mit_Sicherheit_Bauen_04_final.pdf), but again the latter is light

centralized OSH service, making research easier, and the second being the most populous and industrial state in Germany), the website of the Federal Ministry for Labour and Social Affairs[547], and that of the Federal Agency for Labour Protection and Occupational Medicine[548]. In comparison, for Britain, we have only had to peruse the HSE website, which acts as a unique portal for health and safety issues. The conclusions are that the German websites are often inadequate for the needs of business operators, managers and workers, having fragmented information, complex documents, lack of practical examples etc. It is also difficult to know which portal to use. Finally, when real efforts are made to make the information easier to find and to use (as in Nordrhein-Westfalen), this is done at the *Land* level, and thus likely to be ignored in other parts of the country. Each *Land* duplicates the others' efforts, and best results are not shared.

In Britain, by contrast, the HSE website is the single portal for all things "health and safety". One of the site's main tabs is "Guidance[549]" – which then leads into several clear sections e.g. "Industries" and "Topics", which makes information search easy. There is a specific section on "Risk Management"[550] including a set of interactive tools and check-lists for different types of premises. There are clear, detailed and practical brochures for a number of key types of risks, e.g. working with weights[551], "slips and trips" (one of the most frequent causes of accidents on the workplace)[552], or types of establishments such as construction[553]. In the latter page users can, crucially, find a guide to "absolutely essential" health and safety advice in construction, with very practical, clear, visual explanations. This short review leaves no doubt that finding information on OSH issues and good practices is considerably easier in Britain, and that the information is also far more usable. Sources of information in Germany are both dispersed and complex, and the best guidance documents are not available on national (federal) websites and thus are probably ignored in other regions. Of note is also that, while the HSE's efforts in developing and communicating advice and guidance are long standing, they are also part of an increasingly coherent government effort in the UK. In 2009, the Better Regulation Executive published the *Anderson Review* of regulatory guidance[554], which emphasized the importance of making guidance more accessible and clearer – areas in which clearly the available information in Britain is far superior

---

on practical recommendations. The brochure on risk assessment is detailed and practical (though very text-heavy), and is one of the best examples of guidance we have seen on German websites, but the link included in the portal is indirect, and it takes efforts to eventually find and download it (https://broschueren.nordrheinwestfalendirekt.de/herunterladen/der/datei/bro-gefaehrdungsbeurteilung-april2014-pdf/von/gefaehrdungsbeurteilung-am-arbeitsplatz/vom/staatskanzlei/1650).

[547] *Bundesministerium für Arbeit und Soziales* – website accessed at: http://www.bmas.de/DE/Startseite/start.html - pages on labour safety issues e.g. http://www.bmas.de/DE/Themen/Arbeitsschutz/inhalt.html, and for legislation in this area see e.g. http://www.bmas.de/DE/Service/Gesetze/arbstaettv.html. On OSH, the website has primarily (a) general descriptions of issues and policy activities in different areas and (b) federal legislation. Practical guidance, if any, is very limited. There is also no easy collection of links to other institutions, federal and state-level. Hence this website is not really usable as an OSH portal.

[548] *Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA)* (portal at: http://www.baua.de/). The portal is heavy on research papers and legislation, but also has guidance documents on topics that can be very useful, e.g. on handling heavy weights – but these documents tend to be very complex and break down the assessment and recommendations in many components. E.g. there are two brochures to cover the "heavy weitghts" – lifting and handling: http://www.baua.de/de/Themen-von-A-Z/Physische-Belastung/pdf/LMM-Heben-Halten-Tragen.pdf?__blob=publicationFile&v=3, pushing and pulling: http://www.baua.de/de/Themen-von-A-Z/Physische-Belastung/pdf/LMM-Ziehen-Schieben.pdf?__blob=publicationFile&v=3. There is a very detailed brochure also on handwork positions: http://www.baua.de/de/Themen-von-A-Z/Physische-Belastung/pdf/LMM-Manuelle-Arbeit.pdf?__blob=publicationFile&v=9. Overall, information is fragmented, and far from easy to use. It targets rather specialists than operators.

[549] Available at: http://www.hse.gov.uk/guidance/index.htm - see for "topics" http://www.hse.gov.uk/guidance/topics.htm and for "industries" http://www.hse.gov.uk/guidance/industries.htm

[550] See: http://www.hse.gov.uk/risk/index.htm - and tools at: http://www.hse.gov.uk/pubns/indg163.htm (guidance) and http://www.hse.gov.uk/risk/risk-assessment-and-policy-template.doc (template)

[551] See: http://www.hse.gov.uk/pubns/indg143.pdf (brochure with clear, visual guidance), http://www.hse.gov.uk/pubns/indg383.pdf (detailed version with assessment charts), http://www.hse.gov.uk/pubns/indg398.pdf (guidance on use of handling aids).

[552] See: http://www.hse.gov.uk/slips/index.htm including a number of assessment tools, practical tips and brochures etc.

[553] See: http://www.hse.gov.uk/construction/index.htm and "absolutely essential" health and safety brochure at: http://www.hse.gov.uk/pubns/indg344.pdf

[554] Available at: http://webarchive.nationalarchives.gov.uk/20090609003228/http://www.berr.gov.uk/files/file49881.pdf

to that in Germany. The Review also recommended to improve *certainty* and we will briefly discuss this aspect in this section's conclusion.

The gap in effectiveness between Germany and Britain could also have parts of its origin in the earlier development and adoption of a risk-based approach in OSH in Britain. Following the 1974 Health and Safety Act, risk assessment became the norm. In Britain, OSH risk was thus focused on identifying and preventing or mitigating risks, driving efforts to identify trends and evolutions, concentrate resources, provide guidance on the problems most commonly identified (including on problems affecting many different sectors). Rather than prescribing standards, the 1974 Act emphasized outcomes and left flexibility on how best to achieve them. By contrast, in Germany, OSH remained essentially based on sector-specific, detailed standards-based regulations. The 1989 EC Directive on health and safety at work (89/391/EEC), which introduced "general principles of prevention" based on risk (and on risk-assessment and risk-management) (cf. Tilindyte 2012, pp. 92-96) led to significant changes in the German legal order (whereas it ended up requiring remarkably little changes in Britain, as the Directive to a significant extent drew on the same principles as the 1974 Act)[555]. The Directive was translated in German law with the adoption of the 1996 Occupational Health and Safety Act (see Tilindyte 2012 pp. 165-166). While the current system (in particular the mandatory insurers' requirements and practices) is still largely based on pre-1996 principles, including detailed sector-specific rules, and while the uptake of risk assessment by businesses has been slow and difficult (*ibid.* p. 183), there is little doubt that this legal change had important consequences. Still, it is plausible that the more than 20 years gap in implementing risk-based approaches in OSH regulations may be one of the reasons for the gap in efficiency and effectiveness.

The fragmentation of the German system goes together with what appears to be a somewhat weaker management of information, and a less formalized risk-based planning system. German authorities, in fact, used to have regular, "individualized" supervision of businesses largely independently of risk (Tilindyte 2012 p. 182), before the reduction in resources and change in approach led to more targeting (p. 183). As Tilindyte notes, however, much of the targeting is done "in the absence of *formal* models or comprehensive risk assessment models" (*ibid.*), with some targeting based on sectors, other on issues, and many cases simply on the inspectors' "personal experience and expertise". The inspectors' previously very broad autonomy is gradually reducing, but formalized risk assessment models exist so far only in some *Länder* (*ibid.*, p. 184). By contrast, risk-assessment and risk-based targeting are far more clearly formulated and more strongly and consistently implemented in OSH inspections in Britain (though not without room for improvement) – and this is evidently a decisive factor in the fact that Britain manages to achieve excellent OSH outcomes (comparing to EU average for instance) with far less inspections than Germany.

OSH inspections in Britain are explicitly planned on a risk-assessment basis[556], and meaningful risk-assessment requires both *data* and an *information management system.* On this count, the dual structure of OSH enforcement (HSE and LAs) leads to a sub-optimal structure. HSE records inspection activity and findings (and enforcement follow-up) using a database called COIN (Corporate Operational Information System)[557]. Every intervention is recorded in the system and the findings result in an "inspection rating" – taking into account "past performance as well as demonstrated attitudes towards health and safety" (Tilindyte 2012, p. 125). These inspection ratings are then combined with the sector-based prioritization to determine which premises

[555] This is of course a very much shortened account of the process, which was considerably more contentious. The Directive, though using the language of risk, imposes mandatory prescriptions and does not use any language comparable to the British "So Far As Reasonably Practicable" (SFARP). In fact, the "SFARP" clauses were challenged before the CJEU in 2005 by the European Commission, but the CJEU found in favour of the UK (Tilindyte 2012 p. 108). As a result, the British legal framework for OSH was little affected by this Directive (though further, specific directives e.g. on protective equipment have of course been taken up, but they belong more to the "conformity assessment" field than to the OSH field in terms of their implementation and enforcement mechanisms).

[556] See further for a discussion of the changes in terms of targeting introduced by Government policy in recent years, in particular the notion that there could be "high risk" activities but where inspections have been demonstrated to be ineffective.

[557] See on the HSE website: http://www.hse.gov.uk/statistics/data-quality-statement.htm and also Tilindyte (2012) p. 125

will actually be targeted. Significant efforts are made to ensure the integrity and quality of the COIN database, however recent reports suggest that it may not be fully adequate to ensure optimal targeting. In particular, in the 2014 research report conducted by the Health and Safety Laboratory as part of the review of the new "Fee For Intervention" scheme (see also below on this), inspectors reported significant issues with COIN data: "challenges for both [proactive and reactive visits] related to the quality of information available on COIN. For proactive visits, a major challenge discussed was that the intelligence provided was not targeting a sufficient number of organisations with poor health and safety standards" (p. 2). The report further details that "challenges associated with the preparation for proactive visits related to the quality of site information that was available and, more broadly, the intelligence provided to inspectors. In particular, inspectors expressed frustration with the availability of company information on COIN. It was described that very often company information (such as previous inspections or investigations) is not available on COIN because details get deleted for companies that have not been inspected for seven years" (p. 28). It appears that the HSE's targeting system is hampered here by a double challenge. First, inspections are becoming more rare because of decreases in staffing and of a political drive to reduce their number, which means that inspectors are more rarely visiting premises, and thus that updated information on establishments under supervision becomes rare. Second, the quality of targeting suffers from a well-known problem affecting organizations that moved early to computerized data: problems linked to legacy systems. In this case, the information system (COIN) is set up so that data older than seven years is deleted, which means that prior track record will be lost (for companies that have been visited more recently), and that other establishments see all their data disappear altogether from the database (if they have not been visited for the last seven years).

As for LAs, each one uses its own IT system, and they only provide consolidated summary data to the HSE each year (over 90% of LAs effectively do it). While the lack of data sharing between LAs and HSE is in principle not a problem (since supervision and enforcement are clearly divided between them, and there is no establishment where both could be involved in OSH supervision), the fragmentation of data among different LAs is an issue for companies that operate in a number of localities. Efforts to address this problem exist, e.g. as part of the *Primary Authority* scheme (see further below), but this remains a real limitation.

Overall, in spite of these limitations, there is significant evidence that greater emphasis on use of data (in particular records of previous inspections) and clear criteria for risk-based targeting are instrumental in making Britain's OSH inspections more efficient than German ones, i.e. better able to cover key risks for a given number of visits.

Finally, and though there is at this stage no evidence of this link (other than correlation), it is plausible that the way enforcement is structured, and discretion is framed, has some influence on the results. Indeed, we have seen that in Britain OSH inspectors' discretion is exercised within a particular set of institutions, practices and cultures (Hawkins 2002) but also, increasingly, within an "enforcement management framework" that gives more transparency and predictability, and explicitly links the exercise of discretion to an assessment of risk. Post-enforcement discretion of course exists in Germany (Tilindyte 2012 pp. 187-188), within the limits of Administrative Procedures Law (*Verwaltungsverfahrengesetz*), and depending on the wording of the law being enforced (much like in Britain and elsewhere) – but there are no specific guidelines to help inspectors take decisions, and make these more transparent. There are a number of general principles, notably "equal treatment", "proportionality", "necessity" etc. (*ibid.,* p. 188). All of these, however, require interpretation, and leave considerably more leeway and uncertainty than the very specific guidance included in the HSE's Enforcement Management Model (EMM). There are two ways in which the clarity, predictability and transparency introduced by the EMM could have positive effects in terms of safety. First, it could plausibly have a positive procedural justice effect, by making it clearer for duty holders how they will be assessed, and how decisions will be taken – resulting in an increased likelihood of voluntary compliance. Second, because the EMM emphasizes the importance of *risk assessment* and of the "risk gap", it may push businesses to focus

on critical issues and concentrate their compliance efforts on the points that are likely to have the greatest safety impact. In other words, if the EMM is understood (at least in general) by businesses, and if they can reliably expect that enforcement decisions will be risk-based, they will have a strong incentive to focus on improving the points that may cause the most harm and thus (assuming limited resources and inevitably imperfect compliance) maximize the safety outcomes for a given "percentage" of compliance. By contrast, German businesses, left with far more uncertainty and far less clarity of what inspectors expect, are more likely to try and increase compliance in all directions, or focus on points that appear clear but may have less impact. At this point, we do not have any specific data that could prove these relationships, but this would be an interesting area to explore in future research.

iv. Areas of potential concern

We have seen above that there is very strong data suggesting that Britain's OSH inspections system and approach are significantly more efficient than Germany's (comparable safety outcomes with several times less inspections), and that this appears to be linked to a more consistent use of *risk* in planning and decision-making, and more emphasis on *guidance and advice*, as well as a system that is generally more *structured.* This does not mean, however, that all is perfect in the British system and that there are no areas of concern. First, we have seen that Britain's OSH performance *used to be* significantly *better* than Germany's, but that the latter has partly caught up (because Britain's has not continued to improve) – although the trend is less clear when traffic-related accidents (which are not supervised by the HSE) are excluded. Second, we have noted some concerns related to the quality of data used for risk-based targeting. We have also discussed above the many critics of the latest years' evolutions. We have also indicated that there seemed to be serious under-performance in the area of work-related transport accidents. We will first discuss the "data quality" issue (and also respond to observations made in Germany's case), then consider the question of transport-related accidents, and conclude by considering the most recent reforms and their likely effects on OSH effectiveness in Britain.

Data quality, IT systems and low frequency of visits

A relatively frequently made claim is that introducing a comprehensive database of establishments, including their risk profiles, is far too resource-intensive and costly, and this is often used to resist pushes to introduce risk-based planning. Tilindyte (2012) thus writes (in respect to Germany) that "clearly, the effort necessary to establish, maintain and implement such models is substantial" (p. 185), and goes on to emphasize the number of workplaces supervised in each *Land.* Like Baldwin (2007), she goes on to suggest that the costs "to collect, to process and evaluate information on all of them" are a "mammoth task", imposing "high costs" – and that the necessary regular updates will again imply further high costs (p. 185). As we have already noted in the theoretical discussion on risk-based targeting, there is reason to believe that such concerns are overblown (and, possibly, not expressed in an entirely candid way). In reality, as we have seen, German inspectors (if we consider both mandatory insurers and *Land* inspectors) visit nearly 400,000 enterprises each year. Granted, this is only a bit more than 10% of the total number of businesses, but this is more than all the enterprises having 10 employees or more. In other words, German inspectors, assuming that they pool their efforts and all (mandatory insurers and *Gewerbeaufsicht*) enter data in a unified system, could cover all the most "meaningful" enterprises in one year, and all the enterprises with at least one employee in at most a couple of years, simply by conducting their existing number of visits. We have observed the experimental verification of such an undertaking's feasibility in Mongolia. In 2010, the World Bank Group took a group of officials from the General Agency for Specialized Inspections (GASI) to Bosnia and Herzegovina, and they observed the

effectiveness of the unified database and information management system in place in the State Inspectorate of the *Republika Srpska* entity. Over the next couple of years, GASI's management instructed all inspectors to henceforth conduct a risk assessment during every inspection visit, and to enter the data from this risk assessment in a computerized database[558]. After only a couple of years, GASI built up a comprehensive data of all establishments, with their risk profile, and it is being regularly updated. It appears difficult to believe that German inspectors could not do the same – but it would require the setup of a unified system for OSH inspections, and also a clear decision to focus on premises with significant risk (i.e. premises with at least 10 employees at first, moving on afterwards to premises with more than 5, and eventually to those with at least 1 employee – and essentially leaving out those with no employees).

The situation is different in Britain – better for now, but with the perspective for getting worse. Indeed, in contrast to Germany, OSH inspections are now *really* rare (equivalent to less than half of the enterprises with 10 employees or more, and a small fraction of those with at least one employee). Because computerized databases already exist for HSE and most LAs, the problem is one of *updating* – and of *maintenance* for the HSE system, given that premises with no inspection in the last 7 years get "reset" to zero. As we have indicated above, there is a dual information technology problem (a "legacy" system in HSE, fragmented systems in LAs) – combined with an operational challenge created by the decrease in the number of inspection visits overall. It is not the purpose of this research to develop detailed responses to the challenges identified, but this one deserves consideration for, if it was impossible to address it, this could seriously undermine the viability of a risk-based approach. In fact, even assuming that inspections remain at a sustained low rate as is currently the case, there would be practical means to get more regular updates on the risk profile of establishments. This would involve the replacement of current legacy systems (at national and local levels) by a new *integrated* system with a common database for several types of inspections (involving several national agencies and all the different regulatory areas covered by LAs) or, at a minimum, a system enabling regular data sharing between the different existing systems and databases. We have discussed above the existing models and systems that exemplify such approaches (cf. also World Bank Group 2014b and Blanc 2012, as well as OECD 2015 b), and there is no doubt that such data sharing is feasible with existing technology. The challenges involved in Britain would be institutional and political, particularly when it comes to building a system connecting national and local institutions. So far, significant efforts at data sharing between different inspection areas have been done in some local authorities, but broader integration would involve significant steps politically and administratively. In addition, while the idea that other agencies can act as "eyes and ears", that compliance problems in one area can often be predictors of issues in another areas, that the fundamental characteristics of an establishment can be assessed by one inspector for the benefit of several agencies are all correct, it may still be important to reconsider the current trend towards ever-decreasing HSE inspections.

Transport-related accidents

We have seen above that the fatal accident rate reported by HSE, i.e. *excluding* transportation-related accidents, has Britain as the EU's best performer, and with a far lower rate than Germany. When considering the rate of fatal accidents *including* transportation-related "at work" accidents, Britain still features among the "best in EU", but not anymore quite at the top, and not better than Germany, at least in recent years. The fact that HSE reports generally the rate *excluding* transportation accidents has been criticized (e.g. by Tombs and Whyte 2010) as leading to under-estimation of the seriousness of OSH problems, and indeed it tends to distort reality and present a picture more flattering than should be the case. The source of the issue, in fact, appears to be regulatory provisions rather than a specific attempt to understate the issue (and, in fact, this

[558] Sources: interviews with GASI management and GASI central risk assessment unit, review of internal GASI reports for 2012-2014, direct consultation of the database in September 2014.

regulatory requirement much predates the reforms of the past few years). Under the Reporting of Injuries, Diseases and Dangerous Occurrences Regulations (RIDDOR) which govern occupational accidents reporting, work-related accidents on public roads are *not* reportable to HSE, for responsibility for investigating them lies with the police. HSE would only get involved if the police identify any serious management failings that they feel led to the accident and as a result refer it to the HSE – and such referrals are, in practice, quite rare[559]. However, even when work-related transport is taking place on public roads, health and safety law still applies and employers are under a duty to manage occupational road risk in the same way as any other risk – and HSE and LAs inspectors could perfectly include this risk in their preventive and inspection work.

In practice, it seems the absence of a reporting requirement to HSE has been an incentive for health and safety inspectors to avoid focusing on occupational road risk. This issue has been known for a long time: the Royal Society for the Prevention of Accidents (RoSPA) has been lobbying for a greater focus on road accidents "at work" for many years, and laid out findings and recommendations in a 2002 report. This report estimated that between ¼ and ⅓ of road fatalities in Britain may be work related, and that "at work" road accidents may make up more fatalities than all other occupational injuries (Eurostat data indeed suggests the two to be roughly equivalent, or road accidents possibly a bit more numerous than all other types). It suggested that "car and van drivers who cover 25,000 miles a year as part of their job" were working in the equivalent of "acknowledged high hazard sectors such as construction and quarrying" (p. 1). It found, however, that "health and safety law is not applied on the road" in spite of existing legal duties (*ibid.*). Among the changes considered in the report was making road accidents reportable under RIDDOR (p. 2) and ensuring that employers manage risk on the road as part of health and safety (p. 3). To this aim, the report suggested that HSE provide guidance to employers, to conduct a major awareness campaign, and to link more effectively police and HSE enforcement (p. 3). More than a decade afterwards, however, none of these recommendations has been implemented, and the additional resources the report called for have given way to a sharp reduction in resources.

What is particularly important here is that this is an illustration of the well-known adage "what gets measured gets done". In spite of limited (and shrinking) resources, there is solid evidence that the HSE and LAs in Britain have had very good results in keeping fatal occupational injuries at what is, seen from an international perspective, a very low rate. It seems, however, that they have devoted at best minimal attention (or in many cases probably none at all), and that this has had very real effects in terms of sustained rates of "at work" road accidents that are on par with countries that overall have far worse OSH performance than Britain. In spite of limited resources, which would mean that any investment in preventing road accidents may reduce resources available elsewhere, there is some reason to suspect that at least a moderate investment in awareness and prevention, and the inclusion of road safety issues when reviewing risk management plans, could have a net beneficial impact. Rather than concluding, with Tombs and Whyte, that the non-inclusion of these accidents is deliberate and deceitful, we would rather find that it is the result of poor initial regulatory design (RIDDOR), institutional inertia and the unavoidable effect of incentives – if HSE's performance assessment does *not* include road accidents, then it would not be logical for HSE to invest resources in preventing them (to the possible detriment of other areas where its performance is being assessed). The importance of adequately defining *performance indicators* and the *scope of the risks* that an agency is supposed to work on preventing is thus demonstrated.

"Fee For Intervention" and the reduction of proactive inspections

[559] Author's interview with experienced H&S inspectors in Britain.

We have seen above that the significant decrease in inspections that took place in the past 10 years did not seem to harm Britain's OSH performance in aggregate, but this does not means that some of the evolutions do not present some ground for concern. This is the case in particular with the recent drive to reduce all inspections further and *particularly* proactive one, and to introduce a scheme known as "Fee For Intervention" (FFI). This scheme was introduced in 2012 with the aim to "shift some of the cost of health and safety regulation from the public purse to businesses and organisations that break health and safety laws"[560]. The scheme's principle is that if inspectors, when visiting a business "see material breaches of the law", the duty holder "will have to pay a fee[561]". The introduction of FFI was based on the Government's policy paper *Good Health and Safety, Good for Everyone*[562]. This policy paper had three main components: clamping down on "rogue health and safety advisers", reviewing and simplifying regulations (and providing clear and simple guidance, in particular for SMEs) – and "shift the focus of health and safety activity away from businesses that do the right thing, and concentrate on higher risk areas and on dealing with serious breaches of health and safety regulation". FFI pertained to this last objective[563], and resulted from its further articulation in the following way: "this will mean a very substantial drop in the number of health and safety inspections carried out in the UK. We will also shift the cost burden of health and safety away from the taxpayer, and instead make those organisations that gain competitive advantage by flouting the rules pay for the costs of putting things right". While the objective itself may be uncontroversial (reduce burden on compliant businesses, focus on high risks and "repeat offenders"), the *ways* proposed to achieve it ("substantial drop" in inspections, and shifting of the cost burden) are neither obviously logically connected to the goal, nor necessary to achieve it – and quite possibly may in fact work contrary, at least in part, to this stated objective. To understand this better, we need to consider separately the questions of "cost burden" (FFI), and of the number of inspections.

"Making offenders pay" is one of these mottos that has the appeal of its simplicity. It reminds one of "make polluters pay", and superficially could be thought to rely on a similar economic logic – just as charges for pollution aim at ensuring negative externalities are priced in, charges for OSH offenses could do the same, and make compliance more frequent by tilting the economic incentives in its favour. However, as we have discussed at length in the second chapter, there is at best weak evidence that various forms of deterrence (i.e. tilting economic incentives) are really effective at promoting compliance, and the (unequal, relatively weak) effects of deterrence frequently conflict with opposite effects caused by deterrence strategies harming voluntary (ethics-based) compliance. It is thus far from clear that the introduction of FFI should be expected to have a positive effect on compliance (and, ultimately, safety levels). The level of the fees (GBP 124 per hour charged[564]) is in any case most likely too low to create any significant incentive for large businesses, while it may on the contrary end up being perceived as a major (and unjust) burden if applied to SMEs.

The purpose of FFI's introduction clearly appears to be reducing budget expenditures rather than "fairness" or "effectiveness". Even the report by the "Independent FFI Review Panel" (2014)[565], which was written clearly from a "positive" perspective[566], had as its first conclusion that in spite of "challenges associated with FFI" they could "see no viable alternative to it within the current environment for public expenditure" (p. 2). In other words, considering the Government's budget policy, and the seemingly low priority put on HSE budget by the

---

[560] HSE, *Guidance on the application of Fee for Intervention (FFI)*, 2012 (latest revision 2014) – available at: www.hse.gov.uk/pubns/hse47.htm - p. 1

[561] *Ibid.,* p. 6

[562] 2011 – available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/66745/good-health-and-safety.pdf

[563] Listed in second position in the policy paper, p. 3

[564] See the FFI page on the HSE website: http://www.hse.gov.uk/fee-for-intervention/index.htm

[565] Available at: http://www.hse.gov.uk/fee-for-intervention/independent-ffi-review-panel-final-report-2014.pdf

[566] A comparison of the report text with the appendices shows that the drafters took the conclusions as far into positive territory as possible considering the (limited) evidence collected

Department for Work and Pensions (for which HSE's mandate is far from its core 'business' of managing and delivering benefits), FFI was the only remaining option to fill a major funding gap. If we look at the total fees collected so far under the FFI scheme[567] they amount (for October 2012 to May 2015) to a modest GBP 2 Million, to be compared with a total HSE expenditure in financial year 2014-2015 of over GBP 152 Million (HSE Annual Report 2014-2015). Thus, it seems that the cost/benefit value of the scheme may be very unfavourable, considering the significant time costs of administering it for staff[568], and the very real risk of unintended, negative consequences.

To sum up these, we will briefly quote the independent review report – its very euphemisms pointing out the many problems FFI is still expected to pose, and the authors being keen to note that they have not materialized *yet* (which, after only a year and half, is not surprising): "the consistency of approach adopted by HSE inspectors has minimised the detrimental effects", there is "a cost to pay in terms of the relationship between dutyholders and inspectors, particularly with respect to the advice that inspectors feel able to offer businesses, and that they are motivated to seek from inspectors. But the evidence we have considered suggests that this cost has not been as high as was predicted before the introduction of FFI" and finally that the review "could find no compelling evidence to suggest that HSE is using FFI as a 'cash cow', solely to generate revenue". In other words, the very solid and structured professionalism of HSE, and the strength of the existing relationships with duty holders, minimized the harm done by the scheme to HSE's ability to be effective at promoting compliance through guidance and advice, and the ability of its inspectors to keep their professional judgement unaffected by funding considerations. This does *not* mean that FFI, if extended long enough, would not seriously undermine HSE's professionalism – it only means that it has not done so yet. The limited amount of fees collected, making FFI income a marginal source overall, certainly contributed to this.

FFI is different from previously existing cost-recovery areas of HSE's work (e.g. the "major hazards" work under Control of major accident hazards regulations – COMAH), or from similar risk- or complexity-based fees levied e.g. by the Environment Agency of England and Wales or the Swedish Fire Inspection, because FFI is linked to the discovery of problems, and is thus a quasi-automatic "fine" in case of violations, and one that is proportional not to the seriousness of the issue, but to the time HSE will spend on the case. While there is a safeguard, namely that FFI should only be imposed when there is a "material breach[569]" of health and safety regulations, i.e. a "contravention of health and safety law" that is serious enough to require the inspector to issue a written notice. In a way, this is an automatic fine added to the improvement (or prohibition) notice, but one that is proportional only to time spent, and not to risk or harm (or to undue profit), thus being at odds with the overall HSE enforcement approach. Because it is specifically linked to time spent, it endangers the focus of HSE on advice, and the readiness of duty holders to ask for and receive such advice.

The risks posed by FFI may seem minor, considering the review's results and the HSE's well established practices and professional standards. Still, it is important to keep in mind that the effects of such perverse incentives as built in by the FFI scheme work over time, and can lead to quite dramatic results. We have already discussed above the terrible effects caused by a number of municipal police forces in the United States relying primarily on fines and penalties for funding, resulting in police actions that aimed not at securing citizens and maximizing compliance, but at maximizing recorded and penalized violations. We have observed closely a number of agencies that were allowed to keep a part of the fines they issued (or other mandatory payments they imposed, e.g. testing), and in every case it led to a clear worsening of practices, creating strong incentives

[567] Available on the FFI page of the HSE website – direct link: http://www.hse.gov.uk/fee-for-intervention/ffi-invoice-oct12-may15.pdf
[568] See Appendix 1 to the Review Panel Report, available at: http://www.hse.gov.uk/fee-for-intervention/ffi-review-appendix-1.pdf
[569] See HSE, *Guidance on the application of Fee for Intervention (FFI)*, 2012 (latest revision 2014) – available at: www.hse.gov.uk/pubns/hse47.htm - p. 8

to maximize the number of violations found[570]. The fact that the Government imposed at the same time a target to reduce the overall number of inspections means that FFI cannot (in such a context) result in the HSE trying to maximize the number of inspections, but it distorts the relation with duty holders, and the incentives for inspectors, with very few benefits (even financial ones being marginal).

The other element of the new strategy that was meant to "reduce burden" on compliant businesses was the reduction in proactive inspections. This is one more decision which has a very weak link with the stated goal, and only can appear logical if considered very superficially. It is, in fact, very problematic. First, as we have discussed extensively above, *proactive* inspections are in fact widely considered to be the more effective, more efficient kind, as opposed to *reactive* ones which (as Tilindyte 2012 noted) have for them to be the "low-cost" option (since they require no targeting and no intelligence gathering and rely on accidents, complaints or findings of previous visits). A proper risk-based approach can, by definition, *not* be primarily based on reactive inspections, as we have noted. In fact, HSE constantly conducted far more proactive inspections than reactive ones over all the years we were able to consider, and this was a fundamental part of its risk-based approach. While the ratio remains clearly in favour of proactive inspections in HSE (which are 3 to 4 times as numerous as reactive ones in the latest available annual report for 2014-2015), this Government priority on reducing proactive inspections has led to their collapse in LAs inspections (but as we noted above, the "other" category ballooned, suggesting that many visits were just "re-labelled"). Reactive inspections may be "justified" in a "punitive" vision of inspections, but they are simply not the most effective from a risk-based prevention perspective[571].

A valid question would be whether reducing OSH inspections in Britain, at this stage, makes any sense – from a burden perspective, and from an effectiveness and efficiency one. We have seen above that OSH inspections in Britain are several times more rare than in Germany, and also far below the levels found in France. We have also noted that this may be reaching (or be close to reaching) the point where too rare visits make it difficult to maintain a robust risk-based targeting system (though there may be solutions in sharing data with other regulators). How, then, do health and safety inspections compare to *other types of inspections* within Britain, when they are thus singled out for reduction? In order to do so, we looked at the data compiled as part of the "Focus on Enforcement" initiative of the Department for Business Innovation and Skills[572] - with the latest available year being 2012-2013. Four agencies[573] stand out as making up the bulk of "business inspections": the Environment Agency (EA), the HSE, the Food Standards Agency (FSA) and the Animal Health and Veterinary Laboratory Agency (AHVLA). Unfortunately, whereas the FSA reported the inspections made by local authorities on its behalf, the EA did not (HSE also did not, but we have this data from their website, and we have already presented it. Thus, we can only say that the EA inspections are slightly less frequent than the HSE's (around 25,000 per year), but the EA uses another regulatory instrument far more often (permits, with around 500,000 permit holders). It can be expected that LAs environment inspections add a significant tally to

[570] This was true of course in countries and agencies with major corruption problems (e.g. Ukraine's Tax Service and State Committee for Consumer Protection and Standardization in 2006-2009, Tajikistan's Tax Inspectorate in 2004-2007 etc.) but also in agencies with a real commitment to reform and making significant efforts to reduce corruption and increase effectiveness (e.g. Mongolia's GASI in 2014-2015). In all cases, funding incentives meant that inspectors tended to inspect more frequently than needed from a risk perspective, and to find as many violations as possible (and impose fines in every case).

[571] Again for comparison, the General Agency for Specialized Inspections (GASI) in Mongolia has been making progress in introducing risk-based targeting, but its inspectors keep being distracted from it by the practice of following up on every single complaint through an inspection, resulting in over 60% of inspections being reactive. There is no doubt that it results in major inefficiencies. The HSE of course has solid criteria to filter out complaints, and only conducts reactive inspections in cases worthy of investigation, but the general idea that proactive inspections should be reduced by executive fiat is clearly *not* in line with risk-based principles.

[572] See Focus on Enforcement website at: http://discuss.bis.gov.uk/focusonenforcement/ and the link to the data for 2012-2013 at: http://discuss.bis.gov.uk/focusonenforcement/regulator-data-201213/

[573] The initiative did not cover Her Majesty's Revenue and Customs (HMRC), which of course conducts a significant number of inspections.

this number. But OSH inspections are dwarfed by the number of food safety visits: more than 540,000 for the FSA, and over 100,000 for AHVLA. Most of these visits were in fact conducted by third-parties and reported by the national regulator (for the FSA, these third-parties are LAs). Thus, in 2012-2013, OSH inspections made up less than 140,000, and broadly "food-safety related" ones over 640,000. In this perspective, it is not clear that targeting a reduction of OSH inspections made sense even from a pure burden-reduction angle.

It may be that food safety issues warrant a higher level of control from a risk perspective (this would require detailed research on the prevalence of OSH and food risks in Britain, which is beyond our scope), but it may also be that food safety inspection numbers are "shielded" from reduction by a combination of popular perceptions of risk (food safety being usually high among public concerns), EU regulations, and the fact that nearly all inspections are done by LAs. By contrast, it appears that "health and safety" gets targeted for reductions because of a combination of mistaken beliefs ("health and safety myths"), lack of strong interest in the topic from the HSE's parent department (Work and Pensions), and possibly the vagueness of the term ("health and safety" is so broad that many people may attribute to HSE regulations and controls that are completely unrelated to its activities). Whichever the causes, in any case, it seems reducing what is already clearly a very "lean" inspection number is unlikely to have positive results. As we have discussed in the second chapter, there is some evidence suggesting that there may be a lower "threshold" under which the perceived probability of detection is so low that violations *do* increase (breakdown in deterrence). While there is no evidence that this has been reached yet, this pitfall should be taken in consideration for the future.

v. Conclusion – Risk-based and "smart" inspections in Britain

From the above, we can conclude with some confidence that OSH inspections in Britain are indeed a valid example of risk-based, risk-proportionate inspections (both in terms of planning and of enforcement decisions). Indeed, the ways in which the HSE in particular emphasizes guidance and compliance support appears to make this a real example of "smart inspections" in the full sense of the word: risk-focused in inspections, risk-proportionate, transparent, aiming at promoting compliance rather than maximizing "outputs" such as inspection visits, violations identified etc. The comparison of both outputs (inspection visits) and outcomes (rate of fatal occupational injuries) with Germany strongly suggests that Britain's OSH systems is considerably more efficient (at least similar outcomes, with far less inspections) – and even that it was, at least until a couple of years ago, apparently more effective (significantly better outcomes).

While attributing these outcomes fully to the inspections and enforcement system is impossible, and we cannot thus exclude that Britain's better performance is largely or partly due to other factors (still to be specified), there is reason to think that at least some significant part of this higher efficiency is due to "smarter inspections" – and, before that, to a much earlier regulatory focus on *risk*. As we have seen, the emphasis on risk assessment and risk management in OSH dates back, for Britain, at least to the 1974 Act, and translated into very different practices not only for inspectors, but for businesses. A similar evolution, as yet unfinished, did not start in Germany before the 1996 Law translating the 1989 EC directive into German legislation.

At the same time, we have seen that this does not mean that Britain's OSH inspections regime is "optimal" from a risk and "smartness" perspective. First, because of some limitations in the quality and interconnection of data used for targeting. Second, because some of the evolutions in recent years (e.g. FFI) are causes for concern, and may be running *against* a sound risk-based approach rather than being (as they purport to be) its continuation. Third, of course, because perfection in such matters is never possible, and there will inevitably be areas where practices could be improved – the non-inclusion of road accidents "at work" being one of the most obvious areas for improvement.

To put these findings in the perspective of the broader inspections and enforcement system in Britain, it is worth looking briefly at some interesting practices and observations in another regulatory sphere, food safety, which we have selected along with OSH as one of our areas of focus – and also at recent changes that affect inspections and enforcement in all regulatory areas.

"Safer Food, Better Business" and Food Hygiene Ratings

We have noted briefly above that food-related inspections in Britain are far more numerous than OSH inspections. It is difficult, without considerable further research, to draw any conclusion from this in terms of whether this means that risks are significantly higher in food, and/or that "high-risk" premises are more numerous – or whether it reflects a lower level of focus and an approach that is more "inspections heavy". There may be elements of both, considering the large number of food handling premises where risks are not negligible (hospitality and catering), and at the same time the existence of EU legislation that mandates a significant level of inspections (e.g. in slaughterhouses). Rather, we will consider two important examples of "smart approaches" – the "Safer Food, Better Business" toolkit, and the use of Food Hygiene Ratings.

"Safer Food, Better Business"

The 2009 Anderson review stated that Government guidance to businesses should be clear, accessible and consistent. While the review looked at health and safety and employment law guidance, and we have seen that the HSE has made significant efforts in this direction, one of the most interesting examples of guidance is in the food safety field. The origins of the "Safer Food, Better Business" (SFBB) toolkit[574] are to be found in practical experience, and feedback from inspectors on what they found. These resulted in the development of a toolkit that essentially "translates" all essential requirements for catering businesses (including all the EU Hygiene Package provisions, and thus including – crucially – the "HACCP" approach) in a way that is readily understandable by cooks and other employees. The SFBB guide (which exists as a printed pack, PDF or online tool) makes all guidance visual, explains the logic of requirements, and structures them in categories that correspond to the fundamental dimensions of food safety in the kitchen (cross-contamination, cleaning, chilling, cooking, management). It includes a diary (refillable) to keep all mandatory records.

The approach taken in SFBB can be traced back to the finding that many catering businesses had fundamental problems with compliance because of ignorance or misunderstanding of safety requirements, and that this required an approach based on guidance and compliance promotion, including outreach to the many professionals working in the UK but speaking a foreign language. One of the experiments leading to the development of SFBB was made in Chinatown by the Westminster City authorities[575]. After finding that non-compliances in restaurants were not only frequent, but not improving after repeated inspections, the Westminster regulatory team attempted to understand why. They found out that chefs mostly did not really understand English well, were not aware of local safety regulations, changed repeatedly, and that an inspection with negative findings resulted in a loss of face that made compliance, if anything, even less likely. The response was to emphasize prior training, and to use the chefs' language as much as possible. In consequence, the SFBB toolkit exists in 16 languages, those most widespread among chefs working in the UK.

The development and launch of SFBB came in response to the entry into force of the EU "Hygiene Package" and of Government concerns that compliance could prove very difficult for SMEs and in particular small catering businesses – and that this difficulty could come more from difficulties in understanding the

[574] See the portal: http://www.food.gov.uk/business-industry/caterers/sfbb
[575] Short case study: http://www.cieh.org/library/Knowledge/Food_safety_and_hygiene/Case_studies/Westminster%20CHIP.pdf

requirements, than from material challenges[576]. Both national and local evaluations showed strongly positive results after the first year of the scheme already. First, "perceptions of external assistance were very positive. The FSA information packs were very well received with the vast majority wanting to see these continue in future. There were equally positive about local authority interventions, particularly one to one coaching and training courses" (Jigsaw Research 2007, p. 4). The implementation of Food Safety Management appears to have strongly benefited from SFBB. From an evaluation's survey of food business operators, "based on their own perception, 93% of businesses claimed to have FSM in place, and of those, around three-quarters claimed it was fully implemented" (*ibid.,* p.3). Direct evidence from local authorities' inspections provided somewhat lower numbers (unsurprisingly), but still clear confirmation of positive effects: "Compliance with the food safety management requirement for all food businesses has improved from 30% in 2002 to 45% in March 2006 to 48% in March 2007. These figures are based on local authority inspections. Many businesses that have been helped using SFBB have yet to be re-inspected (as the average inspection frequency of the target businesses is once every 18 months). The 48% figure for compliance in March 2007 therefore under-reports the effect of the programme moving businesses towards full compliance. In addition, many businesses are close to compliance and actively working toward it. Evidence from the local authority grant projects shows that where support has been provided, 66% of businesses are broadly compliant shortly after the intervention in the business" (FSA 2007, p. 3).

Assessing and evaluating more precisely the degree to which SFBB has improved (or not) food safety levels in Britain compared to other EU countries, and the extent to which it has helped food businesses (and particularly small ones) make the transition to the new EU legislation, would both require considerable efforts, in particular given the vast number of other factors affecting both indicators. What is important is the strong emphasis put by the British authorities[577] on this approach, and the scarcity of similar examples elsewhere. The only similar guidelines we could find in the EU were developed in the region of Lombardy (Italy) and published in 2014[578]. This Italian document (*Manuale di buone pratiche di igiene per le microimprese alimentari* – "Good Practice Handbook for Hygiene in Food Micro-Businesses") is very interesting in that it covers also food processing, and is adapted to the specifics of Italian food, but it is somewhat less "granular" than the SFBB toolkit and does not as conveniently break down the control steps and points. Still, it remains a very interesting document, that is unfortunately sub-optimally publicized (it is available on a variety of regional and

---

[576] See Hogg (2007): "In response to the new Food Hygiene regulations (EC Regulation 852/2004 on the hygiene of foodstuffs) for the UK, which came into effect on 01 January 2006, the Food Standards Agency (FSA) introduced a new food safety management system based on the principles of HACCP (hazard analysis critical control point) called Safer Food Better Business (SFBB). This system is much simpler than the traditional HACCP methodology in that it cuts out all of the jargon and can be tailored to meet each individual food business needs throughout the country. It is also a very simple system for food businesses to put in place." (p. 4) – available at: http://www.torridge.gov.uk/CHttpHandler.ashx?id=265&p=0 - and from the first official FSA evaluation of the scheme: "The aim of the programme is to help micro1 and small catering and retail businesses comply with the 2006 food hygiene regulations that require businesses to put in place effective food safety management procedures. There are over 400,000 micro food premises in England. Research has shown these businesses previously found food safety management difficult to implement and that improving their standards would have a positive impact on public health by reducing food poisoning. Given the numbers of premises, the scale of activity represents a significant challenge for the Agency and its partners." (FSA 2007, p. 2 – available at: http://tna.europarchive.org/20120419000433/http://www.food.gov.uk/multimedia/pdfs/board/fsa071204.pdf)

[577] Note: the Food Standards Agency that used to work in all of Britain has now been replaced by a Food Standards Agency for England and Wales, and Food Standards Scotland– because Scotland disagreed with the narrowing down of the FSA mandate by the Coalition Government. Food Standards Scotland has the same emphasis on guidance and risk-based approaches and to all intent and purposes this institutional change does not affect the points we discuss here. Among existing nuances in approach, Scotland has had since 2006 a guidance for implementation of HACCP approaches that is somewhat different from SFBB, though similar in concept, called "CookSafe" – see http://www.foodstandards.gov.scot/cooksafe

[578] The publication date is estimated from the date of publication of the different webpages referencing the guidelines, as well as the dates of local authorities' decisions given for context – see e.g. The page on the site of the Azienda Sanitaria Locale (Monza and Brianza): shttp://www.aslmonzabrianza.it/ita/Default.aspx?SEZ=1&PAG=135&NOT=979. The toolkit can be downloaded here: http://www.aslmonzabrianza.it/user/download.aspx?FILE=OBJ02262.PDF&TIPO=FLE&NOME=manuale_buone_pratiche_microimprese_alimentari_RL or here: http://www.asl.lecco.it/intranet/docs_file/6_6582%20allegato.pdf

local websites in Lombardy, but not on national websites, and it is not featured prominently, contrary to SFBB, meaning that many businesses are probably unaware of it). We were not able to ascertain if this Italian handbook was inspired by SFBB, although this seems highly likely. The SFBB toolkit has proven to be inspiring for national authorities in a number of countries where it was presenting, even though adaptation has been rare so far – beyond the Italian example, one more case is Mongolia, where the World Bank Group has supported the translation and adaptation of SFBB into Mongolian, and its distribution to inspectors and local businesses[579].

Food Hygiene Ratings

Assigning ratings to food businesses based on inspection findings in terms of hygiene is not a new idea, and it does not appear to have originated in Britain. The first example we are aware of, and which is documented in scholarship, is from Los Angeles County in the United States, and was introduced in 1998. An evaluation showed substantial beneficial effects on compliance of this scheme, which also involved a risk-based targeting approach (Fielding *et al.* 2001). Overall, the study found that "the development of risk categories has resulted in better targeting of high-risk establishments through more frequent inspections" and that "Retail food establishments behave as if achieving a high letter grade is sufficiently important to comply more strictly with food safety practices. Of particular note is the almost 75% decrease in establishments scoring below 70%, suggesting the impact not only of greater public access to inspection scores, but also of the revised program's more stringent penalty for establishments scoring below 70 twice in a 12-month period. The decrease in closure rates (…) also suggests improved adherence to public health standards for food handling" (*ibid.,* p. 242). Of particular importance for us here is, beyond the vindication of risk-based targeting and the note on the importance of deterrence for repeat offenders, the effectiveness of publicly available ratings at increasing compliance. Similar schemes were later introduced in other locations, e.g. in Denmark (from 2001), which has long been one of the best known such schemes[580]. Similar schemes have been introduced or discussed in countries as diverse as Nepal and China, Germany and France, Kyrgyzstan[581] etc. Thus, Britain's use of Food Hygiene Ratings is neither an isolated nor a "pioneering" case – but it deserves consideration in how it fits with Britain's approach to food safety inspections and enforcement more broadly.

Food hygiene ratings were introduced much later in Britain than in Denmark: "the FHRS[582] was launched in late 2010 and local authority uptake following that progressed rapidly such that the scheme is now well bedded in. All but one authority (Rutland County Council) in the three countries[583] is now operating the scheme and information is available on the FSA website on over 440,000 food businesses. The FHIS[584] is now operating across Scotland. The FHRS was put on a statutory footing in Wales in November 2013 to provide for mandatory display of FHRS ratings at food premises. The transition from the voluntary scheme to the statutory scheme is due to be complete in May this year. Northern Ireland is set to follow suit with draft legislation introducing mandatory display currently being considered by the Northern Ireland Assembly, and the FSA strategy for 2015 to 2020 highlights that pressing the case to extend this to England is a priority"[585]. As this report indicates, the display of food hygiene ratings is still not mandatory in England, though all businesses do get a rating after

[579] Author's direct interviews with World Bank and Mongolian government officials involved.
[580] See overview of the system here: http://www.findsmiley.dk/en-US/Forside.htm and here http://en.fvm.dk/focus-on/smiley-food-inspection/.
[581] Author's direct observations (press reports, pilot schemes) in France, World Bank colleagues internal reports on Germany and China, author's direct involvement in pilot schemes in Nepal and Kyrgyzstan.
[582] Food Hygiene Ratings Scheme
[583] England, Northern Ireland and Wales.
[584] Food Hygiene Information Scheme
[585] Source: Food Standards Agency Board Meeting – 25 March 2015 (FSA 15/03/06), *Food Hygiene Rating Scheme – Update And Next Steps,* Report by Jason Feeney, Chief Operating Officer – available at: http://www.food.gov.uk/sites/default/files/fsa150306.pdf

each inspection, and consumers can freely access these on line. Furthermore, non-display of the rating is likely to indicate that the rating is low, and the FSA website logically advises caution to consumers in such cases.

There is an important difference between the FHRS (England, Wales, Northern Ireland) and the FHIS (Scotland), which is the number of possible different ratings. The FHRS has zero to five stars[586], whereas the FHIS only has "pass" or "improvement required"[587] (by comparison, the Danish scheme has four types of smileys, plus the "elite smiley" for those repeatedly obtaining a clean sheet). The FSA scheme thus provides more nuanced and detailed information (at the risk of some potential for confusion on the exact meaning of such nuances for consumers), whereas the FSS goes for a very simple system. In any case, what is noteworthy is that *ratings* are distinct from *enforcement* decisions, and introduce an additional tool for compliance promotion, with more flexibility than enforcement measures.

Indeed, as the FSA's "Frequently Asked Questions" on FHRS make clear, even a business with a low rating is not necessarily always shut down (or at least not immediately) – low ratings are not equivalent *ipso facto* to formal enforcement measures, but rather are a "stronger form" of informal enforcement, using the powerful incentive (market effect) created by the ratings' publicity. Quoting the FSA website: "Businesses given ratings of '0' or '1' must make urgent or major improvements to hygiene standards. The local authority food safety officer will use a number of enforcement tools as well as giving advice and guidance to make sure these improvements are made. (…) The Food Hygiene Rating Scheme means that people can choose instead to eat out or buy food at places with higher ratings and businesses with low ratings are in danger of losing customers and so will be encouraged to improve standards more quickly and to maintain these in the future. If the officer finds that a business's hygiene standards are very poor and there is an imminent risk to health – this means food is not safe to eat – the officer must take action to make sure that consumers are protected. This could mean prohibiting part of an operation or closing the business down[588]." This additional, intermediate form of "semi-formal enforcement" can be appropriate in a number of situations where the inspector thinks shutting down operations is inadequate (negative social impact, loss of revenue that will make improvements even more difficult to achieve etc.). The ratings are a tool whereby the administrative action (inspections and enforcement) uses market forces to ensure increases in compliance. They are interesting in that they offer high flexibility (because formal enforcement is not automatic, there is reversibility in the decision if things improve), and businesses can ask for a re-visit if they have made the required improvement. By offering transparency to consumers, less drastic measures than formal enforcement (and a chance for improvement) to businesses, food hygiene ratings thus are an instrument that seems to be able to combine a real "deterrence" effect (through reputational risk) and a high level of procedural justice.

The March 2015 report to the FSA's board includes the summary of an evaluation (conducted over 2011-2014) which shows that (i) "consumer awareness and reported use of the FHRS have steadily increased to 36% and 20% respectively, and 76% of people recognise the distinctive green and black branding", (ii) "there was a significant increase in 'broad compliance' (ratings of 3 to 5) in the first year, and a significant increase in 'full compliance' (rating of 5) in the second year in local authority areas after FHRS was introduced, compared with areas where the scheme was not yet operating" and "there was also a significant decrease in the proportion of very poorly performing businesses in the first two years after launch"[589]. Thus validating the findings of the 2001 survey of the Los Angeles County case, this FSA evaluation shows (including by comparing "treatment"

---

[586] See FSA website at: http://www.food.gov.uk/multimedia/hygiene-rating-schemes/ratings-find-out-more-en/fhrs

[587] See FSS website at: http://www.foodstandards.gov.scot/food-safety-standards/food-safety-hygiene/food-hygiene-information-scheme

[588] "Why are businesses with poor ratings not closed?" in FAQ at: http://www.food.gov.uk/multimedia/hygiene-rating-schemes/ratings-find-out-more-en/fhrs

[589] Source: Food Standards Agency Board Meeting – 25 March 2015 (FSA 15/03/06), *Food Hygiene Rating Scheme – Update And Next Steps,* Report by Jason Feeney, Chief Operating Officer – available at: http://www.food.gov.uk/sites/default/files/fsa150306.pdf - p. 3

and "control" groups given phased introduction of the scheme) that public food hygiene ratings are (gradually at least) used by consumers, and have a positive impact on compliance.

Of course, it is important to remember that such food hygiene rating schemes are only useful if the professionalism and ethics of the inspectors assigning the ratings is such that (a) consumers will have trust in the ratings and pay attention to them (something that, of course, also requires adequate efforts to "advertise" the rating scheme) and (b) they will not abuse the power given by these ratings (since bad ratings can be highly damaging to businesses) for corrupt purposes. Such potential pitfalls are not limited to low- and middle-income countries. The 2001 evaluation of the Los Angeles County case pointed out that "Strong economic incentives to achieve higher grades increases the risk of bribery.(…) One confirmed episode of an inspector soliciting a bribe, captured in an investigative report on a local news station in November 1998, prompted DHS to review and revise measures in place to prevent illegal and unauthorized activities by inspectors" (Fielding *et al.* 2001, p. 244). Credibility and visibility can also be real issues. Some initial pilot schemes in France had little success, and it remains to be seen whether the new initiative to publicize the inspection results (as ratings in four letters) from July 2015 (to be extended country-wide in 2016) will be more successful[590]. Thus, food hygiene rating schemes are not a solution that can be implemented independently from a broader risk-based approach, and require to be successful a strong basis of professionalism among inspectors.

The *Primary Authority* scheme

The United Kingdom (and Britain, within it, in particular) present the specific situation of regulations that are primarily national (or, in some cases, European) with inspections and enforcement that are primarily conducted by officers hired by local self-governing bodies. In a 2009 report, the then-Local Better Regulation Office (LBRO – now BRDO) counted 433 councils administering regulatory enforcement across the UK (out of which 407 in Britain) – to which should be added 58 Fire and Rescue Authorities (57 in Britain) and 151 Port Health Authorities (141 in Britain)[591]. Reviews for a number of years, such as those conducted by Hampton or Löfstedt, have been pointing out the issue of consistency as a key area for improvement. The Regulatory Enforcement and Sanctions Act (2008), apart from establishing the possibility of introducing new types of administrative sanctions (as recommended by the Macrory review) also created the "Primary Authority" statutory scheme, to be administered by LBRO (and, from 2012, by BRDO as its successor body). Though we will only discuss it in a very cursory way, Primary Authority is interesting because it seeks to address several issues: consistency of enforcement between different localities, transparency and clarity of regulations, certainty of advice.

Primary Authority can apply to "a single business that is regulated by multiple local authorities, or to a business that is part of a group of businesses that are collectively regulated by multiple local authorities, where these businesses share an approach to compliance" that "might be demonstrated, for example, through membership of a trade association that provides regulatory guidance, or through a franchisee relationship with a business that specifies compliance controls"[592]. The business(es) enter(s) a partnership with a single local authority ('primary authority') which has to be approved by BRDO, which will validate it only if it has assessed this particular authority as adequately competent in the regulatory area(s) under consideration. The role of the primary authority is then to be a "key point of contact" in relation to the business' interactions with

---

[590] See newspaper reports: http://next.liberation.fr/food/2015/07/01/quel-niveau-d-hygiene-dans-les-restaurants-pres-de-chez-vous_1340907 and http://next.liberation.fr/food/2015/07/01/y-a-t-il-vraiment-un-probleme-d-hygiene-dans-les-restaurants-asiatiques_1313183 - official news release and data here: http://agriculture.gouv.fr/experimentation-de-la-mise-en-transparence-des-controles-officiels-en-restauration-commerciale-paris

[591] LBRO 2009, pp. 25-27

[592] BRDO, *Primary Authority Statutory Guidance,* September 2013 – p. 3

local authorities that regulate it and to coordinate the regulatory enforcement efforts in relation to this business[593]. It does this e.g. by providing "advice and guidance on compliance to the business (known as 'Primary Authority Advice') in areas of regulation covered by the partnership, on which the business can rely" and also by providing advice and guidance to *other local authorities* in regard to this business. It may "publish an inspection plan" to guide and co-ordinate their enforcement activities. Crucially, the advice given by the primary authority to the business can be opposed to another local authority: "where the business faces potential enforcement action by an enforcing authority, the primary authority will assess whether the proposed action is inconsistent with any Primary Authority Advice given. If the action is inconsistent, the primary authority is able to direct the enforcing authority not to take the action". If the primary authority has published an inspection plan, other authorities have to follow it. They also must "notify the primary authority of enforcement action in relation to the business" and "in most circumstances this notification is required before the action can be taken" (except for emergencies)[594]. In other words, the primary authority for a business can give it in-depth advice that, if it then follows, it can be assured will be found to be valid by other authorities – and it establishes a significant degree of consistency and coordination in inspection plans and enforcement responses. Because of the significant amount of work involved for the primary authority, its work within such a partnership is done on a cost-recovery basis.

The benefits for businesses are clear: they get assurance that, if they behave in a certain way, this will be found to be in compliance with the law – and that the way they are treated will be as uniform as possible across the country. There are significant benefits expected for citizens as well (be they consumers, neighbours, workers etc.): if key regulatory requirements are "internalized" in the internal directives and procedures of businesses, they are far more likely to be complied with. Ensuring that advice and guidance are given by the most competent local authorities in a given field is also expected to improve the overall regulatory outcomes. And, of course, higher certainty of regulatory interpretation should lead to some uptick in investment and growth – and lower regulatory burden to (marginally) lower prices.

Primary Authority has been the subject of several evaluations and studies – a first evaluation in 2009-2011 (conducted by RAND Europe), a second in 2013 by acl Consulting, and an evaluation of the impact of training sessions on primary authority for inspectors, by Dunlop, Kamkhaji and Radaelli in 2013-2015. As the scheme is not universally applied (contrary to, say, the food hygiene ratings) and uptake is voluntary, and since it is in many ways a very significant change, it is not surprising that evaluation results were not one-sided. The 2011 review found that positive effects on consistency were not reported by more than 22% of businesses, but many noted that it was "too early", and for most participants it had not been a major concern or primary reason to enter the scheme anyway. The evaluation noted differences between a more "thorough" primary authority relationship (with an inspection plan) and "lighter" ones: when an inspection plan was present, consistency was more strongly increased, and satisfaction levels were higher (in any case ¾ of businesses were satisfied). The study was not designed to really capture compliance levels, but noted that a significant share of local authorities had changed the way they dealt with businesses in case of problems, far more often going to the head office rather than branches (pp. 43-47). The 2013 evaluation found again ¾ of businesses satisfied, but only 45% of local authorities considering the overall impact positive (versus 30% negative and 25% neutral). Businesses generally noted a reduction in time spent on regulatory issues, better relationships with regulators, more consistency in advice and guidance. On the compliance side, it was found to reduce "instances where action is necessary in respect of non-compliance by promoting informal discussions between primary authorities and enforcing authorities" – but the authors noted that most participants in the scheme were already businesses with a "positive approach" and "positive interest" in compliance, making it difficult

[593] *Ibid.*
[594] *Ibid.,* pp. 3-4

to see a clear impact on that front (pp. 5-7). As for the study by Dunlop *et al.* (2015), it did not look at the scheme itself, but at the trainings provided by BRDO to inspectors, and found that they had a significant impact on the extent to which inspectors understood the logic and the relationships with the scheme, and thus their ability to make the most of it (p. 3). Because both evaluations (2011 and 2013) noted that there remained many implementation and understanding issues, this finding on the effectiveness of training is important.

Overall, Primary Authority is a very important innovation conceptually, and there is a large amount of anecdotal evidence of its benefits[595], but there is not yet a set of very strong evaluation findings on its effects. In any case, from our perspective, it demonstrates the importance of considering inspections and enforcement issues comprehensively, not only in terms of targeting or numbers, but of methods, consistency, relations with duty holders, guidance etc. Primary Authority is not yet old enough (and, maybe, not yet widespread enough in use) to make strong conclusions, but does appear to be an approach that has the potential to bring improvements in terms of compliance, interactions between inspectors and enterprises, and overall procedural justice.

Summary conclusion

As we can see from the above OSH case study, and from the snippets presented from inspections and enforcement practices in other areas, Britain (and, more broadly, the UK) can be taken as a valid example of risk-based approaches and, at least in some areas, of "smarter approaches". This does not mean, of course, that inspection practices in OSH or other areas in Britain are "perfectly risk-based" (if such a thing were possible). There are areas where significant improvement from this perspective would clearly be possible, for instance data quality in OSH, or the "twin peaks" created by the distribution of premises between HSE and LAs. In food safety, SFBB and food hygiene ratings seem to be very promising initiatives, but their impact remains to be more fully evaluated, as should be the frequency of inspections. Primary Authority is innovative and seems to have the potential to solve major contradictions and issues in inspections and enforcement, but is still in its growth phase. In addition, institutions, methods and practices are all vulnerable to changes driven by short-term budgetary considerations more than by sound evidence, and the introduction of FFI and sharp reduction of proactive inspections in the OSH area are causes for serious concern, and seem likely to lead to a worsening rather than an improvement of the efficiency and effectiveness of the system.

Still, when considering the overall performance of Britain's OSH system compared to Germany's, the magnitude of the efficiency gap is such that it strongly suggests that Britain's risk-based approach seems to have real benefits. With several times less frequent inspections, Britain has consistently performed at least as well as (and, until a few years ago, significantly better than) Germany. This is certainly not a result that is due entirely to different inspection practices (and, for instance, the far earlier emphasis on risk-management *within enterprises* in Britain probably also played a role), it also seems unlikely that Britain's far more risk-focused approach is not one of the reasons for this difference in performance. Investigating whether such greater efficiency and positive results can be seen in other areas (e.g. food safety) would thus be a very valid area of future research.

b. Post-Soviet and Post-Communist Experiences

[595] See BRDO (2014), *Primary Authority: A Guide for Officials*, available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/348664/14-1058-pa-guide-for-officials.pdf

The "post-Soviet" or "former Soviet" world (or, to use a convenient moniker, the Former Soviet Union or FSU[596]) continues, in spite of the more than 20 years elapsed since the USSR's collapse, and of the obvious deep differences between its constituents, to share a number of common characteristics, particularly in the structures and workings of the public administration. This is true also, albeit to a lesser and more unequal extent, for countries that were never part of the Soviet Union, but were under its direct influence (Comecon members)[597]. Decades of shared institutions, legal systems, practices and experience (longer for some, shorter for others) resulted in regulatory systems that were, at least at the onset, very close. Of course, various backgrounds in terms of earlier history and social structure, economy, culture and geography meant there were major differences between all these countries – but also considerable similarities, among which was precisely the way economic activity was organized (and, at it gradually became private, regulated).

These countries' shared history and structures did not end all of a sudden in 1989-1991, but rather in many cases have effects that can still be easily observed. Nor were privatization (with the many challenges it represented, and the many ways in which it went socially, economically and politically "wrong") or "shock therapy" (where it was attempted) enough to comprehensively transform the relations between state administration and economic operators[598]. In fact, at the end of the 1990s, business regulation, and as part of it inspections and enforcement, were a significant issue in practically all FSU countries, and most if not all other former Communist countries (including some outside of the group we are considering, e.g. former Yugoslav republics). As a result, reforming business regulations and specifically, in a number of cases, business inspections and enforcement, started to become an important priority – both for national governments, international organizations and, in the case of candidate countries, for the EU.

Problems with regulatory systems stemmed from a number of aspects of the Communist-era inheritance. A hostile attitude to private ownership and private businesses fostered confrontational approaches geared at punishing violations that were expected to be numerous, with businesses seen as criminal by essence. The way the law was for decades systematically made to mean whatever was in favour of the authorities left a very weak rule of law and highly problematic ethics in the state administration (and outside of it too), and courts that tended to rule in favour of the most powerful (and, in particular, in favour of the state). Institutions left over from the previous period were strongly specialized, heavily staffed, with a strong technical bend, and a risk-averse approach, resulting in a tendency to try and achieve total control over each and every risk, and also in each risk dimension being controlled repeatedly from several different angles. Of course, this common background was nuanced by country specifics. The Baltic States, for instance, drew on a more liberal tradition, and on the legislative traditions from the inter-war period. Russia, by contrast, had a dreaded figure of 'inspector' long before Soviet times, as exemplified in Gogol's *Inspector General* (in Russian "*Revizor*"). In most of these countries, however, and definitely in all former Soviet ones, the term "inspection" (in Russian

[596] This includes the following fifteen internationally recognized independent countries: Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan

[597] This includes the following: Bulgaria, Czech Republic, Hungary, Mongolia, Poland, Romania, Slovakia. Former Yugoslav Republics had a completely different legal and administrative system, and so had Albania (though the latter was far more "statist" than the former). Vietnam could also be considered, as well as Cuba, but they were considerably more "remote" from the Soviet Union, and present major specificities. We will thus limit ourselves to the seven listed above. East Germany was incorporated into the Federal Republic of Germany and took over "wholesale" its regulatory structures. Some differences may linger in practices of various *Bundesländer* but studying them is beyond the scope of this research.

[598] See e.g. EBRD *Transition Report* for 1999 and 2000. These reports, however, focused on a very limited set of "liberalisation" indicators (e.g. price and quantitative import/export controls) and, while they occasionally mentioned broader "market access" issues, did not really investigate them. In the first decade of "transition", the importance of technical, safety etc. regulations (and their enforcement) was not always perceived, with more "fundamental" aspects of a market economy being still in the forefront.

"*proverka*") and its derivatives were well known by all, and elicited prompt responses from any interviewed citizens, business operator or not[599].

While European integration ensured that, to a large extent, underlying regulations were transformed in candidate countries, it left regulatory procedures largely untouched. Indeed, as we will outline in the next section, the level of harmonization of "regulatory delivery" is overall relatively weak in the EU, even in areas where regulations themselves are partly or strongly harmonized, with the partial exception of food safety and, of market surveillance of non-food products. This means that, in many areas, inspection structures and methods were left largely untouched by the EU accession process (while corruption in "street level bureaucracy" did overall decrease as a result of across-the-board institutional reforms, but was clearly not eliminated). While the stronger level of harmonization in food safety meant that inspection structures and practices did change in this area, this did not necessarily lead to changes in other instruments, e.g. licenses and permits, which were able to remain "on top" of EU requirements. Harmonization in conformity assessment for the "New Approach" directives in regard to non-food products was likewise strong, but did not prevent the persistence of pre-reform inspection practices, additional licensing requirements etc. In short, while EU accession did result in significant changes (particularly if comparing new EU Member States with other former Soviet countries in Eastern Europe, e.g. Ukraine), it did not (by far) transform all pre-existing practices[600].

The salience of the problem was evidenced through action taken by governments in Eastern Europe and Central Asia, with or without the support of international organizations – Latvia was one of the very first, with an inspectorate improvement programme supported by the World Bank Group in 1999-2003[601], Armenia passed a law on inspections in 2000, and Russia a law on the protection of business operators during inspections in 2001. Romania attempted to consolidated a number of inspection functions in 2003, and Poland set procedural guarantees for businesses during inspections in 2004[602]. Lithuania embarked significantly later, but in a more ambitious way, in a thorough reform of inspections (2010 onwards)[603]. Evidence of the prevalence of the problem is also evidenced in the responses to the business surveys that the World Bank Group started running in a number of countries, in support of business environment reform programmes (which increasingly included inspections as well as permits and licensing reforms), from the end of the 1990s. Countries covered by these surveys (at various dates) include Azerbaijan, Belarus, Georgia, Kyrgyzstan, Mongolia, Tajikistan, Ukraine and Uzbekistan[604]. All of them showed, at least before reform (and often after several years of it, in spite of some improvements), the prevalence of inspections-related problems.

The wealth of data collected by these surveys (as well as by some country-specific ones, e.g. those run by the Lithuanian government since 2010)[605] will be a very important source for the outline of inspection practices,

---

[599] See for instance Coolidge, Grava and Putnina 2003 on the situation in Latvia in the late 1990s.

[600] For an example of how such regulatory practices can survive in a new EU Member State, see e.g. OECD 2015 b on Lithuania. The whole report is sub-titled "focus on the delivery side", emphasizing the importance of regulatory procedures and processes rather than only the text of regulations. See in particular chapters 7 (inspections and enforcement) and 9 (construction permits). Lithuania also has a large number of other licenses and permits, including approval requirements for food business operators that are stricter than mandated by EU regulations (source: direct interviews with and presentations by senior officials). Several other OECD reviews of regulatory reforms covered post-Communist countries (Czech Republic, Russia, Kazakhstan) – see: http://www.oecd-ilibrary.org/governance/oecd-reviews-of-regulatory-reform_19900481.

[601] See Coolidge, Grava and Putnina 2003 as well as Putnina 2005 and Coolidge, Grava and Liepina 2008.

[602] See Putnina 2005 and World Bank Group 2010.

[603] See OECD 2015 b, chapter 7.

[604] The surveys covered the following years: Azerbaijan: 2008 (a later survey remained unpublished) – Belarus: 2003, 2004, 2005 and 2008 – Georgia: 2003 and 2005 - Kyrgyzstan: 2008 and 2010 – Mongolia: 2009 and 2015 (forthcoming) - Tajikistan: 2003, 2005 and 2007 (a later survey remained unpublished) – Ukraine: yearly from 2000 to 2004, then 2006, 2008 and 2010 – Uzbekistan: yearly from 2001 to 2007. All these surveys see bibliography IFC (various years).

[605] See OECD chapter 7 for a summary of these surveys. Other surveys covering businesses' experiences in the region are those run jointly by the EBRD and World Bank Group and known collectively as BEEPS (Business Environment and Enterprise Performance

effects and changes in former Communist countries that we will briefly draw below. The other key source will be our own experience – being involved as an advisor to governments in a number of these countries since 2004[606], which resulted in countless discussions and interviews with both government officials, businesses and external experts.

i. Inspections pre-reform: burden and attempt at "total control"

Geographically, this section focuses on several countries of the Former Soviet Union (Armenia, Azerbaijan, Belarus, Georgia, Kyrgyzstan, Tajikistan, Ukraine and Uzbekistan), to which Mongolia (which was very close administratively to the Soviet system) is added. We also consider one EU Member State that is also at the same time a former Soviet republic (Lithuania). There are several reasons to consider this particular set of cases. First, the Former Soviet Union, and more broadly former command-economies of the Soviet block, have been (for many still are) characterized by rigid and heavy regulatory systems which aim at preventing "all risks, all the time, everywhere" – in effect, a direct opposite of the risk-based approaches we are considering in this research. Second, most of these countries have undertaken significant regulatory reforms (often focusing specifically on inspections, and with a risk angle) in the past ten or even fifteen years, and many have reached important results, but also faced significant limitations – thus presenting interesting lessons on how such reforms can succeed (or not), and with what effects. In all these countries, business inspections used to be, and often still are, a major problem, and thus have been a key area of reform – which allows to investigate specifically this aspect, and the impact of changes. Because of the salience of these issues, and of the need to have reliable data to design, steer and evaluate reforms, the International Finance Corporation of the World Bank Group, and the World Bank Group Investment Climate Advisory Services, have conducted business surveys in most of these countries (and the Lithuanian government has done likewise, and with similar questions and methods), focusing on regulatory procedures and instruments, in particular inspections, which provide a wealth of data.

Before reform, the regulatory approach in these countries was (and still is in part in many cases) extremely prescriptive, with detailed "specification-type" rules setting out exactly what material should be on the walls, how a shop or factory should be laid out, what recipe to use to preserve cucumbers, how many coat-hangers should be in a hotel room etc. Not only were these rules highly prescriptive and detailed, but they were hugely numerous, and came from a large number of different regulatory bodies, without coordination (and quite often with contradictions). In order to ensure adherence to these norms, and to control business activity more broadly, most activities were (and often still are) subject to *ex ante* controls: businesses and citizens require hundreds of permits, approvals, licenses etc., which must frequently be renewed. In addition, once operating, businesses are subject to numerous inspections regardless of the actual risk level of activities, and likewise customs, traffic police etc. attempt to control each and every person, truck, shipment.

Evidently, most of these countries (or, probably, all of them, but to varying extents) present clear links between "petty corruption" (that of "street level bureaucrats" rather than major top-level corruption involving large contracts) and frequent inspections (and permits, licenses etc.). This link is not only seen in the circle of

---

Subject). We discuss later in this research issues with the reliability of these surveys, but in any case they essentially do not cover inspections (except in a very marginal way) and thus are of limited relevance to our work.

[606] Working in this period in Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Lithuania, Mongolia, Tajikistan, Ukraine and Uzbekistan. It should be added that, before that, we also experienced first hand government inspections in Tajikistan as a non-profit entity, while working as Central Asia coordinator for the French NGO ACTED. In 2003, a shipment of WHO-recommended insecticide to be used for malaria preventionwas delayed over a month at the border for spurious technical reasons (and very real corrupt demands), until it got escalated by the funding party (the EC) and cleared at the highest political level. This resulted in serious harm (the insecticide only arrived after the start of the malaria transmission season), and definitely drove home to us the importance of the issue.

more inspections resulting in more bribes resulting in continued frequent inspections. It is also observed in that it appears that burdensome and intrusive controls, unfair treatment and constant inspections breed disrespect for the law on the side of the controlled persons and entities. Thus, it appears that even in cases where individuals are not corrupted, these practices tend to "corrupt" the system, i.e. make it inefficient and ineffective. The relevance of these (tentative) findings may be challenged by arguing that the regimes of Central Asia, Russia or Ukraine, are "inherently" authoritarian, plagued by 'cultural issues'. Many (over more than a century) have argued that authoritarian rule is endemic to the 'Russian character'. Others may suggest those governments' are unable to manage modern administrative systems because of lack of resources. This is too easy a way to dismiss the significance of post-Soviet (post-Communist) cases.

Clearly, lack of resources is not the reason. Russia's, Kazakhstan's, Azerbaijan's or Mongolia's GDP per capita, thanks to natural resources wealth, increased rapidly during most of the 2000s (even though Mongolia has slowed down, and Russia and Kazakhstan suffered from drops in oil prices and exchange rates)[607], and are on par with at least some European countries. Belarus (even though this is partly due to its industry still benefiting from a *de facto* energy subsidy from Russia) also has higher PPP GDP per capita than the poorest EU Member States. Even though other countries in our study group are mostly poorer (with the exception of Lithuania), it is sufficient to show that lack of resources cannot be the explanatory variable – particularly given that corruption problems in regulatory issues are definitely not lower in the resource-rich group (except for Mongolia, which has a significantly more open and more democratic system). Nor can the old fallacy of "cultural specificities" explain away the "control and corruption nexus". This fallacy is in fact used most frequently by corrupt senior officials themselves to justify the lack of reforms. In the 19th century, Germans were derided by Britons as lazy and corrupt – and laziness was also one of the main Japanese characteristics according to Western observers in the early 20th century[608]. Nowadays it is held that the German and Japanese cultures embody hard work and that corruption is low in Germany because it is abhorrent to the national culture.

Cultural differences exist, and the way the Soviet Union applied its laws (as tools to root out dissidents, meaning the law would always be against you no matter what it appeared to say) is definitely a key factor in the corruption problem. But one should look closer: most Russians, Ukrainians, Tajiks etc. complain about corruption and see it as a problem[609]. Most officials, for their part, go to great lengths to appear to comply with the laws. In many cases at least, based on numerous conversations with businesses in all these countries, officials generally make references to applicable norms and highlight violations that are at least partly credible, rather than outright asking for a bribe (though at the lowest level, with micro-businesses e.g. in Tajikistan, directly asking rather than bothering with pointing out norms is frequently observed). Most people, both citizens, businesses and officials, clearly see corruption as an evil – even if it is widely practised, even if they

---

[607] In 2014, Russia and Kazakhstan both had higher nominal GDP per capita than Romania and Bulgaria (the poorest EU Member States), and only 10% lower than Hungary. Mongolia's was much lower but catching up already with Tunisia and Albania, for instance. At PPP, Russia was even above Poland and around 10% below Slovakia, Kazakhstan was above Slovakia and Mongolia less than 10% below Serbia
(Source: World Bank – see:
http://data.worldbank.org/indicator/NY.GDP.PCAP.CD?order=wbapi_data_value_2014+wbapi_data_value+wbapi_data_value-last&sort=desc and
http://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD/countries?order=wbapi_data_value_2014+wbapi_data_value+wbapi_data_value-last&sort=desc).

[608] See Chang (2007), Chapter 9

[609] Even though surveys as BEEPS sometimes report the percentage of respondents rating corruption as a "major" or "serious" problems to be decreasing, this is neither because of a decrease in the phenomenon, nor (based on considerable evidence, including successful revolutionary movements in Kyrgyzstan and Ukraine, and unsuccessful movements elsewhere) because of growing acceptance of it, but simply because of the way such questions tend to elicit unreliable responses because of the way respondents may perceive "problems" as meaning "unusual or somewhat solvable issues", as opposed to endemic and permanent ones. See for this our section discussing data reliability issues in surveys such as BEEPS.

benefit from it themselves. They may be profiting from corruption (on any side) but they are not advocating the practice as being "good" or even "normal". Crucially, and by contrast, there is a wide consensus about the way the system *should* work: there should be *effective total control*. Citizens long for it, businesses assume this is how things should work and inspectors are partly pretending, partly genuinely trying to enforce rules (taking bribes is not necessarily in contradiction with thinking one is trying to enforce laws, and sometimes actually enforcing them). Our argument is that this *objective of total control* is precisely a core element of the problem, and one that is therefore worthy of reform – and of study.

All the countries that we are considering here have attempted, with varying levels of commitment, different approaches and unequal success, to reform their inspection systems. These reforms were driven by the recognition that there were blatant problems with existing practices, and they also were part of an international context, where reforms in some countries of the region inspired others, and where the World Bank Group (and, to a somewhat smaller extent, other international assistance actors such as USAID, for instance) promoted such reforms and worked on trying to spread risk-based approaches. Crucially for our purpose, in most of the countries where it implemented programmes to support inspections reform, the World Bank Group ran representative business surveys to capture the frequency, duration and a number of other aspects of inspections (and of a number of other administrative procedures). While there were some variations in sampling structure and in precise wording of questions, these surveys offer a sufficient level of homogeneity to make meaningful comparisons on inspections incidence and frequency, and their evolution, both between the different countries surveyed, and over time. Given the large sample sizes used for most of these surveys (1 to 2 thousand respondents in general), and the relatively high engagement of World Bank Group teams in these countries in quality control, the reliability of data is rather high, at least when one considers simple "objective" questions such as the number of inspection visits. Unfortunately, no equivalent data exists to track the regulatory outcomes such as food safety or occupational safety (for reasons that we will briefly discuss), hence a full comparison of "before" and "after" reform will be impossible, but this will still enable us to get a first impression of the impact of changes.

ii. Data perspective: reform results and international comparisons

How much a regulatory agency inspects is a fundamental metric – be it relative to its staff's other tasks, or from the inspected establishments' perspective (what percentage of them are inspected every year, and how often on average). Data on the share of resources and staff-time spent on inspecting is mostly lacking. Even in OECD and EU countries, many agencies are loath to release such figures, or simply do not track them. It is even more so in the focus countries for our paper, even though discussions with officials suggest that *most* resources and time are spent on inspecting[610].

All the countries in our "surveyed group" shared initially a high level of inspections "volume", i.e. most businesses[611] (75% to 100%[612]) were inspected[613], usually several times a year. Post-reform data, in a majority

[610] Close to 100% of resources spent on inspecting in FSU. Confidential data from regulatory agencies in some OECD countries suggests that, there, at least 20% is spent on analysis and back-office work (and regulatory work includes not just inspecting, but informing).

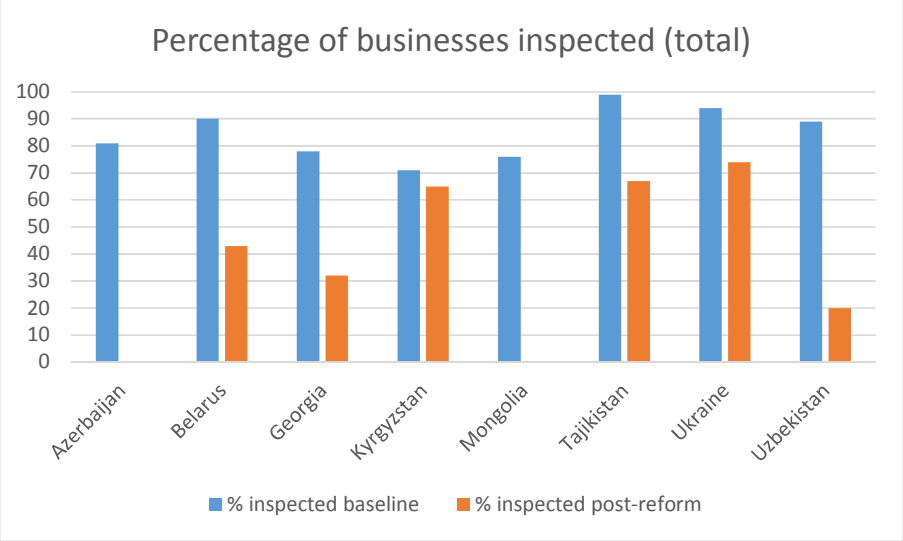
[611] Looking at percentage of *establishments* would be more accurate, as a single business may operate several. In surveyed countries, however, the majority of businesses (and the near-totality of SMEs) correspond to only one establishment. Since, for sampling reasons, surveys were based on *business* population [because business registries are based on entities, not premises), the two are assumed to be essentially equivalent here.

[612] The populations surveyed are not entirely identical, due to differences in the registriation of sole proprietors (and the possibility, or not, to combine their sample with legal entities'), the inclusion or not of agricultural producers, etc. Nonetheless, the general picture is comparable. See survey reports for detailed methodologies.

[613] To assess *targeting*, it would be better to have data on the percentage inspected *out of the supervised population*, i.e. the establishments that the regulator effectively has competence upon, but in most cases, for the countries considered, this population can be equated with the general business population, as regulators have very broad mandates, and make full use of them.

of countries where it is available (i.e. where reform has progressed for long enough, and where new surveys have been conducted), shows a significant decrease. The extent of this decrease, however, varies considerably depending on the character of the reform (more or less radical and/or well implemented, with Georgia the most, and Tajikistan, Kyrgyzstan and Ukraine the least), and on its duration. It appears also that strong, authoritarian governments can (when they are set on doing so) achieve significant decreases in inspections incidence, frequency and duration more easily, in some instances, than relatively weaker or more democratic regimes. These strong decreases, however, often do not reflect profound improvements in the environment for business creation and growth, or real changes in practices, but rather formal compliance with orders from the top. Thus, changes in the percentage of businesses inspected (incidence), the number of annual visits (frequency) or the duration are interesting indicators – but far from the whole story.

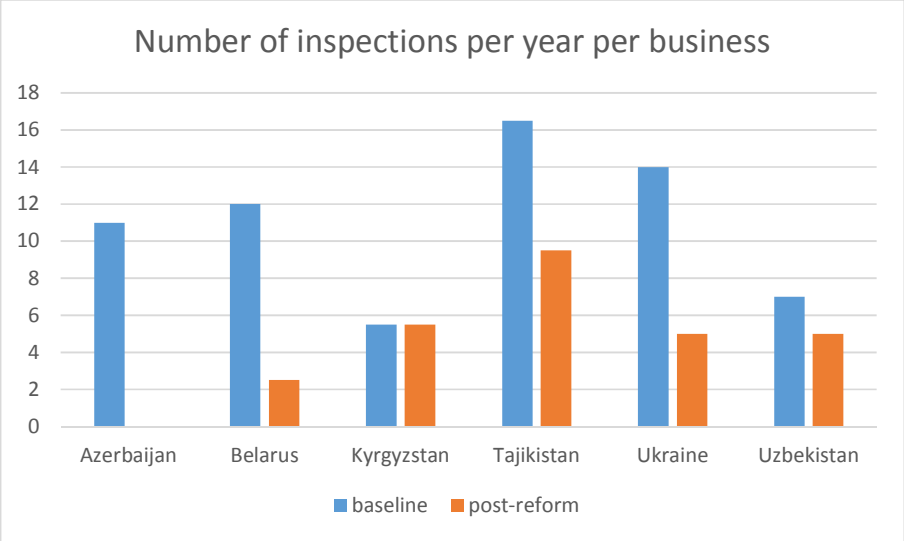
The graph below gives a general picture – with "baseline" and "post-reform" corresponding to different years for each country[614] [note that reform continued in Georgia, and inspections volumes went down much further, but no subsequent survey was conducted, hence the figure below is the latest available]. It shows both the very high incidence of inspections all across our group of countries pre-reform, and the significant decreases in most cases.

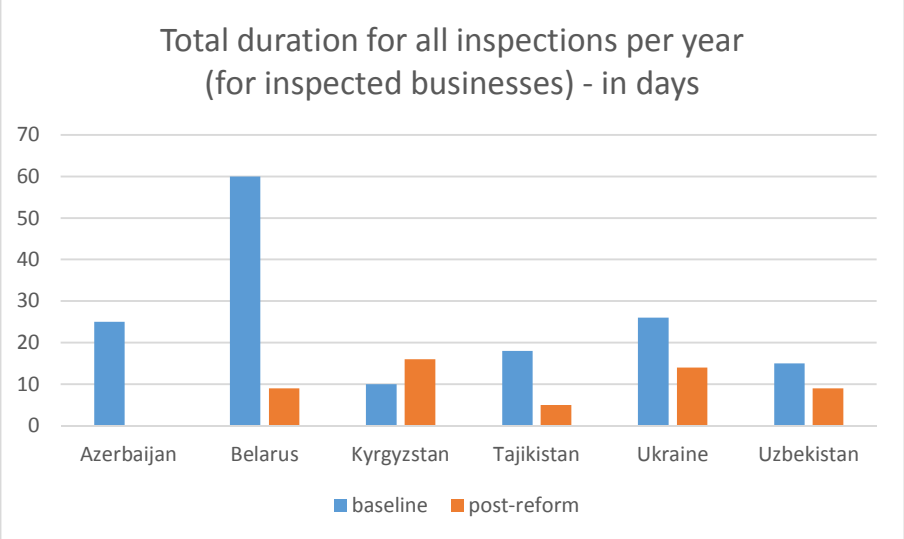


Percentage of businesses inspected in a given year, by country

Incidence is, of course, not the only relevant indicator. Other data, such as the number of inspection visits and their aggregate duration, generally give the same picture, with the same countries displaying strong (or weak) performance, and generally high baselines (though with significant differences) in all cases.

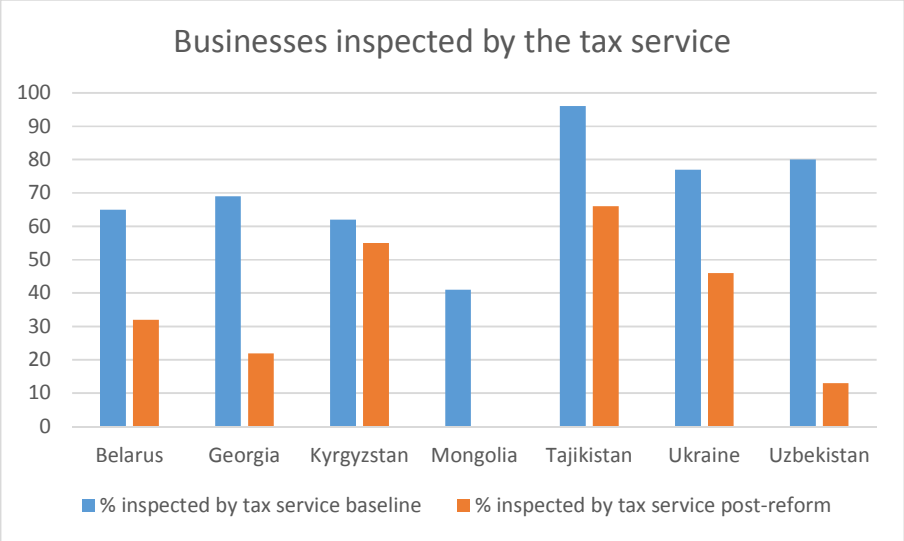
[614] Azerbaijan: 2007 [a subsequent survey was planned but did not take place], Belarus: 2003 and 2012, Georgia: 2003 and 2005, Kyrgyzstan: 2008 and 2011, Mongolia: 2008-9 (Q4 to Q3) [a follow up survey is planned in 2016], Tajikistan: 2003 and 2010 [the 2010 data was presented publicly but not published, it suffers from lower quality than previous years – last published report is 2007], Ukraine: 2006 and 2010, Uzbekistan: 2001 to 2007. Reforms started in: Azerbaijan 2011 only, Belarus 2006, Georgia 2001 with acceleration in 2003, Kyrgyzstan 2005 but stalled several times because of political events, Mongolia 2003 but stalled and resumed in 2009 only, Tajikistan 2006 with slow implementation, Ukraine 2007 with similarly difficult implementation, Uzbekistan 1998 deepened from 2002.

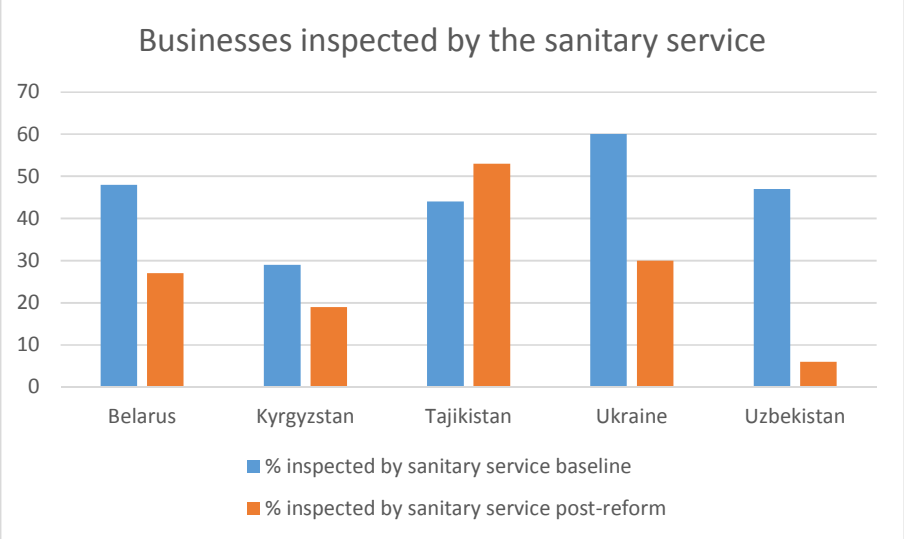




These figures aggregate *all* inspectorates but the bulk of inspections are conducted by one to five (at most ten) agencies, while the remainder[615] are far less active, because of limited staff and resources. Overall, in all countries, tax inspections, fire safety inspections, hygiene and food inspections formed the bulk of the visits. Survey results show the prevalence of these inspections, as well as mixed reform results, corresponding to different degrees of implementation by different agencies.

[615] Just a couple in Mongolia, where most agencies were consolidated in a "single inspectorate" – around 20-30 in Tajikistan, Kyrgyzstan, Uzbekistan – around 80 in Ukraine. This compares with around 70 in Lithuania, similarly with only a few being really active and most of them being very small, or less than 15 in Latvia or Slovenia. It is often difficult to ascertain the number of inspection agencies because many can be small, or local/regional. Thus, while in the Netherlands there is now only a dozen of national inspectorates, an inventory a couple of years ago identified nearly 70 bodies with inspecting powers.





It would be optimal to compare this data to that of EU countries, but it is difficult to do so for, unfortunately, most agencies do not publish data on the percentage of businesses inspected, and very few countries have conducted surveys on this issue[616].

It is in some cases possible to estimate inspections incidence based on (a) the published number of inspections and (b) the total number of businesses in the country[617], but again it is only possible for those who publish their inspection numbers (a small minority). We have to explain here why we think it possible to use data published by inspecting agencies themselves in this case, whereas for FSU countries we have preferred to rely

[616] The data from these surveys is of a different nature, and scope, from the data used in the function-focused case studies. While in the OSH studies (Britain, Germany, and further down France) we have inspectorates' own data, covering only one regulatory function, here we look at survey data with responses from businesses, covering a number of different functions. In countries where there are reasons to doubt the quality of inspectorates' data (which is frequently the case, even in the OECD, as we illustrate below in the case of France OSH – but is far more the case in a post-Soviet context, for instance), business surveys are a superior alternative. Considering percentages inspected, rather than raw number of inspection visits, allows to make easier comparisons country-to-country, regardless of their size. Finally, the ability to look at the overall prevalence of inspections across all functions allows to give a picture that goes beyond a specific function, and reflects the general situation in the country in terms of inspection practices.

[617] This *over-estimates* the percentage of businesses inspected, since some may have visited twice, and this this data is not usually available. This still allows to get an order of magnitude.

on representative surveys. The reliability of data on inspections numbers provided by inspectorates relies on at least three preconditions: adequate information systems to record the data, high compliance with internal rules by frontline officers, and alignment of incentives for the agency, the inspectors and the management (i.e. absence of incentives to manipulate the data one way or another). These conditions are broadly met in the case of the EU-based examples we use below[618], but not in the FSU countries. There, by contrast, not only are information systems frequently absent (except in tax services, records tend to be only paper-based), and compliance with internal processes is far from assured[619]. Most crucially, while there are strong incentives to conduct as many inspections as possible in practice (inspectors and their managers frequently draw illicit income from this, and some inspectorates are also allowed to keep a percentage of the official fines), there is an incentive to report less, since political authorities generally want to at least have appearances of business friendliness, so would react negatively at the very high numbers of inspections that would be reported. Considering official reports by the French tax service[620], and official data from the Latvia tax service[621], they both appear to inspect less than 5% of all businesses, a sharp contrast with the very high percentages observed in the FSU. As for British Occupational Safety and Health inspections, we saw in the previous section that there are less than 200,000 of them per year, for more than 5 million businesses, resulting in a percentage inspected of less than 5% as well.

Doing comparisons for an entire country (all types of inspections together) requires, however, business survey data, for no EU country currently has consolidated data on all inspections conducted in a given year by all agencies. While some have good quality data, this is far from the case for all of them, many do not publish inspection numbers, and there is in any case no way to avoid double counts and deduct from these numbers an aggregate incidence and frequency. Only very few EU countries have conducted representative surveys that we can use for our purpose. While the UK's National Audit Office and Better Regulatory Delivery Office conduct regular surveys of businesses about regulatory matters (in recent years jointly), these do not directly ask whether a business was inspected in a given year, they do not include tax services, and overall do not provide data that would be comparable[622]. The governments of Italy and Lithuania, by contrast, have both conducted surveys that yield (with some minor *caveats*) directly comparable data.

---

[618] When an inspectorate thinks the inspection numbers may be "too high" or "too low" for some important stakeholders, it usually simply refrains from releasing them in the EU. This is e.g. the case of the British HSE in recent years, which prefers to avoid releasing aggregate figures to avoid criticisms one way or another.

[619] Both because of archaic management practices (heavy on authoritarian approaches, weak on real control) and because of incentives (low salary and high prevalence of corruption), actual compliance is low at all levels.

[620] See: Direction Générale des Finances Publiques (DGFIP), *Rapport d'activité 2014 – cahier statistique,* available at: http://www.economie.gouv.fr/files/files/directions_services/dgfip/Rapport/2014/RA_2014_cahierstats_0107_web.pdf. See page 13: in 2014 there were less than 48,000 on site 'accounting verifications' (tax inspections) as well as less than 70,000 checks of compliance with the public television service contribution. Even assuming (though it is highly unlikely) that all these targeted different businesses (whereas in fact a large overlap is likely), the total number receiving any form of "on site" controls was only around 100,000. In the same year, France had more than 3 million non-agricultural enterprises, and over 500,000 agricultural enterprises, hence somewhere between 2 and 3% of enterprises (at most) received an on-site tax inspection [enterprise population see INSEE website – accessed at: http://www.insee.fr/fr/themes/document.asp?ref_id=if4 and http://www.insee.fr/fr/themes/document.asp?ref_id=T13F172]. By contrast, the number of checks conducted remotely (documentation checks) was far higher – but still covered no more than around 10% of all businesses (even assuming that each individual control measure affected a different enterprise, which is again highly unlikely).

[621] Unpublished case study prepared for the International Finance Corporation of the World Bank Group in 2009, showing that less than 5% of enterprises in Latvia received a tax inspection in the most recent year.

[622] Several EU countries have initiated programmes that could yield a consolidated view of inspections, but so far they cover only some sectors, regions or inspectorates, and never all inspectorates in a country (though a few examples of this exist worldwide, e.g. in countries where most inspectorates have been consolidated in one single agency, like in Bosnia and Herzegovina). In Italy, the Registri Unici dei Controlli cover only the agricultural production and processing sector, and so far only at the regional level (though a nationwide extension has been decided). In the Netherlands, the *InspectieView* programme covers only national inspectorates (and not all of them fully at this stage). For more on this topic, see World Bank Group 2014 d.

In Italy there was (so far) a single survey covering 2011 and whereby over 1,500 respondents were interviewed, covering businesses with 5 to 250 employees. Because very small businesses were excluded, there may be a non-insignificant bias in the results (they may be inspected less frequently and, because of their high number in Italy, this would reduce the mean incidence and frequency). The exclusion of large businesses is likely to have less of an impact given their small number (though the mean incidence and frequency of inspections in these businesses is likely to be high). Though detailed results of the survey were presented in a number of events in 2012 by the *Dipartamento della Funzione Pubblica* in charge of administrative simplification, only some excerpts are available on line[623]. This survey showed that the percentage that had received at least one inspection in 2011 was slightly above 36% - hence, far lower than in FSU countries (even though, by many metrics, the Italian authorities considered the inspection burden to be far too high, in particular because of many duplications in controls, and of the high concentration of inspection visits on a few sectors, which received a far higher burden than the average may suggest[624]).

In Lithuania, the percentage inspected pre-reform was high, though somewhat lower than FSU countries that have not joined the EU, at around 60% in 2011 (see OECD 2015 b, p. 122). Initial gains (decrease to less than 45% in 2013) have been partly reversed due to a lower focus on reform under a new administration. If one considers not only incidence but also the combined frequency of visits and duration, the overall inspections burden fell around 30% in the first years of reform and, even after the partial reversal in recent years, this burden remains around 20% lower that at the baseline (*ibid.*, p. 124).

iii. Assessing outcomes of inspection systems – and of reforms

Ideally, in order to properly assess the effectiveness of inspection systems in FSU countries, and of reforms, we would take an approach similar to the one we used to compare British and German OSH inspections, using publicly available data on key outcomes such as food-borne diseases, occupational health etc. Unfortunately, such data is generally highly unreliable in most FSU countries, partly for the same reasons that official data on inspection numbers is not reliable. First, information systems are often lacking entirely, or are limited in scope and usage. Second, political priorities may make under- or over-reporting a much better strategy for managers than giving accurate data. In addition, challenges such as detection (e.g. for food-borne diseases) that apply in all countries are even stronger in countries where the public health system has been in upheaval for a couple of decades and is frequently plagued by corruption. Thus, it is impossible to look for precise correlations at the level of one inspection field, except possibly in taxation, where data on tax income (and on tax income as percent of GDP, and on collections vs. plans) are significantly more reliable (though not perfect).

Being able to compare data on food safety would be very valuable, for it is an area where, in spite of major differences in conditions (natural, economic), changes in practices can make a very strong difference, and one could expect better inspections and enforcement practices to deliver positive results (at least over some years, by supporting changes in how business operators and consumers behave). Unfortunately, except in cases of major outbreaks, food-borne diseases tend to be under-reported: not all patients will see a doctor, and few doctors will prescribe tests to identify the pathogen, except if the case is particularly serious. To this must be added very different practices in terms of health care (in some countries, most patients will go to a general practitioner, if they consult at all – in others, many will go to hospital), different reporting rules and standards. In addition, the consolidation of data from local sources is, in some countries, very problematic – with cases of "sloppiness", but also cases of outright manipulation of data in order to either show better performance

[623] See: *Dossier – I Controlli*, Dipartamento della Funzione Pubblica – Ufficio per la Semplificazione Amministrativa – available at: http://www.funzionepubblica.gov.it/media/1023751/dossier_controlli.pdf (see pp. 21-23).

[624] Our estimates based on the published data suggest that, for businesses with more than 10 employees, the mean number of inspection visits for those having been inspected at least once was around 10 per year in 2011, a very high number indeed.

than is the case or, in other circumstances, make the situation look worse to obtain more funds[625]. While food-borne diseases are *in aggregate* a serious issue, and they rank among the foremost public concerns in developed countries, they are rarely among the most salient risks, or gravest epidemics, and are thus not even reported separately in global health statistics compiled in WHO reports[626]. These WHO reports do however include incidence of under-5 mortality due to diarrhoea (not all of which is linked to food-borne diseases, but which can be a somewhat acceptable proxy), and WHO statistical tables (unfortunately with data over 10 years old, from 2004)[627] include estimates of Disability Adjusted Life Years (DALY) loss for diarrhoeal diseases (which is again an acceptable proxy, and this time for the whole population). The WHO also provides statistical tables with somewhat more recent data (2008) for standardized death rates from diarrhoeal diseases[628] - these are somewhat less helpful (effectiveness of health care system will impact the death rate more than the DALY, and some diarrhoeal diseases will rarely be deadly, so DALY is a better reflection of the actual food safety situation), but combining the two sources allows to cross-check if the situation has evolved significantly between 2004 and 2008.

Reviewing the available data on diarrhoeal morbidity and mortality from WHO tables shows that its reliability is far from perfect[629]. This is unsurprising given what we have noted above. More precise data (on specific causes of disease) is, when available at the national level, even less reliable. Detection of even the most prominent causes of disease, such as *salmonella*, is problematic in most countries. To compensate for detection and reporting biases, advanced regulatory systems such as the EU's are underpinned by systematic monitoring programmes in order to assess (based on sampling and testing) the actual prevalence of key contaminations[630]. Since such monitoring is not conducted in a comparable way (neither in terms of sampling, nor of reliability of tests or consolidation) in countries of the Former Soviet Union (excluding, of course, those that have joined the EU), it is impossible to *precisely* assess their food safety levels[631]. Still, combining WHO statistics and other sources (anecdotal evidence, expert reports, audit reports on the food safety system) it is possible to draw some tentative conclusions.

First, several of the countries with very high food safety and hygiene inspection rates (and which have had such high rates for many years) have clearly dismal records in terms of food safety. This is the case e.g. for Tajikistan, Kyrgyzstan or (demonstrating that this cannot be explained only by low income levels) Azerbaijan.

---

[625] We have witnessed first hand all of these problems in the years 2001-2004 in Tajikistan, not for food-borne diseases but for malaria, a disease that is normally higher profile and thus more systematically diagnosed and reported (and where data quality problems are thus likely to be *less salient* than for food-borne diseases). First, an epidemilogical survey showed that there was very considerable under-detection at the local level. Second, review of the hospital-level data and comparison with aggregated national data showed that there was strong under-reporting once consolidated. Third, sharp variations in consolidated levels year-on-year suggested active manipulation of data for fundraising purpose (while keeping the overall under-reporting to protect the Ministry's image).

[626] See WHO database available at: http://apps.who.int/gho/data/node.home and consolidated reports on World Health Statistics (WHO, yearly) available at: http://www.who.int/gho/publications/world_health_statistics/en/

[627] Available at: http://www.who.int/healthinfo/global_burden_disease/estimates_country/en/

[628] *Ibid.*

[629] For instance, Ukraine has DALY loss for diarrhoeal diseases that is only 30% higher than in the UK and othe EU countries (2004) and even reports age-standardized death rates for diarrhoeal diseases (2008) that are significantly *lower* than e.g. the UK, France or Germany. While this is not *impossible*, other available evidence does not suggest this to be true, but more likely the result of under-detection and/or under-reporting, as the food safety situation in Ukraine is widely agreed to be worse than in these countries (of course, some of this could translate in longer-term health problems rather than diarrhoea, or less easily categorized symptoms, e.g. in case of chemical contamination). Among the "high incidence" countries, there are also very sharp variations between otherwise relatively comparable countries, that all point to the need for caution in using these statistics. They are, however, quite helpful for a first approximation.

[630] See in particular EFSA's monitoring programmes at: http://www.efsa.europa.eu/fr/topics/topic/monitoringandanalysisoffood-bornediseases and summary reports at: http://www.efsa.europa.eu/fr/node/952441

[631] National-level data, when it exists, should be treated with great caution. While working in Ukraine, in 2008, we heard a very senior official of the country's Sanitary and Epidemiological Service deny any problems with food safety regulations in the country by stating that their *salmonella* prevalence was lower than Switzerland's or Norway's. It may very well be that the *officially reported* prevalence was lower, but it was obvious to all that this did not reflect in any way the real situation in the country.

All three have high (>900 per 100,000 for Kyrgyzstan till >1,900 for Tajikistan) DALY loss from diarrhoeal disease (2004) as well as high age-standardized death rates from these same diseases (>9 per 100,000 for Azerbaijan till >26 for Tajikistan – 2008). The US Food Safety and Inspection System has not conducted any audit of any of these countries[632], Tajikistan has never been audited by the EU either, Kyrgyzstan was only audited in 1998 for horses (and the findings were negative, i.e. exports to the EU were not authorized)[633]. Azerbaijan received a few audits, focused on fishery products (2002, 2007) and equidae (2009)[634]. The audits' conclusions on the food safety system as a whole were negative, even though some establishments were found satisfactory. Of course, the EU FVO reports (or their absence) reflect not only on the food safety regulatory system, but on private sector capacity. Very poor countries, landlocked, like Tajikistan and Kyrgyzstan, also simply little that could be of interest to the EU, and no establishments that could meet its requirements – but the negative findings on Azerbaijan show the lack of connection between frequency of inspections and effectiveness. This is confirmed by the many difficulties Kyrgyzstan and Tajikistan have faced in recent years for their exports to Russia and Kazakhstan because of food safety concerns[635].

Further evidence of the disconnection between heavy inspections regimes and effectiveness of the food safety regime is provided by the examples of Georgia and Ukraine. As indicated above, Georgia dismantled most of its regulatory agencies and drastically curtailed regulatory inspections following the "Rose Revolution". With EU assistance, and as a requirement for the conclusion of the Deep and Comprehensive Free Trade Agreement with the EU[636], Georgia has gradually built up a new National Food Agency (NFA). However, in keeping with the market-friendly orientations of government policy, as of 2014[637] the NFA was still conducting only very few inspections (at most a couple of percent of all businesses were being checked yearly). In spite of this "light touch" approach, successive EU FVO reports showed strong improvements (in the fisheries sector, which was the one for which approval for exports was requested) between 2010 and 2014[638]. As for Ukraine, it has received many successive EU FVO audit visits over the year, being a major agricultural producer and food processor, neighbouring the EU, and with a major export potential. While until 2007-2008 most reports highlighted very serious deficiencies, there have been marked improvements in the past 8 to 10 years, with increasing market access in particular for certain Ukrainian eggs and dairy producers (though it is worth noting that this applies only to a few of the best firms in the country, and most of the sector is still excluded, reflecting unequal private sector capacity). In spite of inspections that remain overall very frequent[639], important reforms took place in the years 2008-2014 (and are being continued), in particular the consolidation of all food processing inspections under one agency (State Veterinary and Phytosanitary Service of Ukraine, formed and

---

[632] See on the USDA website: http://www.fsis.usda.gov/wps/portal/fsis/topics/international-affairs/importing-products/2000-2003-foreign-audit-reports and http://www.fsis.usda.gov/wps/portal/fsis/topics/international-affairs/importing-products/eligible-countries-products-foreign-establishments/foreign-audit-reports. The US FDA audit reports are not systematically published and only available on the basis of Freedom of Information Act (FOIA)requests, so this evidence could not be used for this research – see answer on availability of reports on the FDA website: http://www.fda.gov/Food/ComplianceEnforcement/Inspections/ucm211823.htm#q33

[633] See report on the FVO website: http://ec.europa.eu/food/fvo/audit_reports/details.cfm?rep_id=471

[634] There are also 2 reports on aflatoxin contamination in hazelnuts, which are less useful as proxies for the whole food safety system. See reports on the FVO website: http://ec.europa.eu/food/fvo/audit_reports/details.cfm?rep_id=894 - http://ec.europa.eu/food/fvo/audit_reports/details.cfm?rep_id=1894 and - http://ec.europa.eu/food/fvo/audit_reports/details.cfm?rep_id=2378.

[635] Russia is well known for using food-safety based import bans for geopolitical purposes, but the two countries are rather firmly in its sphere of influence and in these cases the bans or threats seemed to at least partly reflect very real concerns. Kazakhstan's measures on meat and dairy from Kyrgyzstan seem to primarily reflect real concerns. Russia's and Kazakhstan's food safety systems are themselves far from perfect, but the overall situation is markedly better than in Tajikistan and Kyrgyzstan, giving some credibility to the assessment. See e.g. several instances of bans on Kyrgyz dairy by Kazakhstan: http://www.akipress.com/news:564986/ and http://en.tengrinews.kz/politics_sub/Kazakhstan-to-lift-ban-on-dairy-products-imports-from-7842/.

[636] This Agreement (DCFTA) was achieved as part of the signature of the Association Agreement in 2014 – see: http://www.eeas.europa.eu/factsheets/news/eu-georgia_factsheet_en.htm

[637] Interview of senior management of the NFA of Georgia by the author, June 2014.

[638]

[639] See World Bank Group (2011), *Investment Climate in Ukraine as Seen by Private Businesses* – pp. 39-42

reorganized in 2010-2012[640]), the abolition of mandatory certification of food products by the State Committee for Standardization, etc. Overall, these changes resulted in a significant decrease in inspections and enforcement (and in other mandatory procedures), even though the post-reform level remains high (average number of inspection visits per year down around 30% between 2006 and 2010 as per World Bank Group survey data). Thus, while it is obvious that Ukraine's inspections system remains closer to the "old post-Soviet" norm than to EU practices, significant improvements in regulatory performance as recorded in successive EU FVO reports[641] took place alongside a significant *decrease* in inspections coverage and frequency.

A further confirmation of the disconnect between inspections "intensity" and public health and safety outcomes is provided by looking at aggregate data for Georgia compared to its neighbours Armenia and Azerbaijan – three countries which, in spite of their differences (in particular very rapid growth in the past 15 years in Azerbaijan due to its massive hydrocarbon resources), share many similarities in terms of starting conditions and level of development. Considering that inspections affect many areas of public safety and health, and that precise indicators are hard to come by and/or present reliability issues, we can take the opposite approach and consider high-level aggregates. Following Helsloot (2012) and Helsloot and Schmidt (2012 a), we can use life-expectancy as a proxy for overall physical safety in the broadest sense, and look at trends in life-expectancy in these three countries. Considering the period before any regulatory reform started (i.e. the late 1990s), life-expectancy at birth was (WHO data[642]) 72 years in Georgia in 2000, 71 years in Armenia and 66 years in Azerbaijan. The same source indicates 74, 71 and 72 years respectively in 2013. World Bank data gives very similar figures[643]. Georgia had a slightly longer life-expectancy at birth before the reforms started, and it still does, with a slight improvement over 2000. Armenia has remained stable. Azerbaijan has experienced a rather strong improvement, which is likely to be primarily linked to the massive increase in wealth over the past 15 years[644] (and possibly the phasing out of some very highly polluting chemical industries inherited from the Soviet period). While Azerbaijan has not done significant reforms in regulations, inspections and enforcement, this is beside the point because their effects would in any case be dwarfed by the increase in incomes (and income is the primary driver of physical safety, cf. Helsloot 2012). What matters to our research is that Georgia, which did very radical reforms resulting in a very sharp drop in inspections and enforcement "intensity" did not see any worsening of life expectancy (and even a small improvement), in spite of the period also being characterized by internal and external warfare, as well as a partial economic embargo imposed by Russia, which heavily weighed on economic recovery. In spite of disaster warnings by some that cutting such inspections would have dramatic safety consequences[645]. Both President Saakashvili and Minister of Economy (and then of Reform Coordination) Bendukidze stated during the reform process that risks were minimal from disbanding these institutions and stopping their activities, because they had hitherto been corrupt and using ineffective approaches. Facts appear to have mostly vindicated them.

---

[640] See EU FVO report on dairy products (2014) available at: http://ec.europa.eu/food/fvo/audit_reports/details.cfm?rep_id=3377

[641] See e.g. meat 2009: http://ec.europa.eu/food/fvo/audit_reports/details.cfm?rep_id=2344 – animal health 2010: http://ec.europa.eu/food/fvo/audit_reports/details.cfm?rep_id=2651 – salmonella in eggs 2013: http://ec.europa.eu/food/fvo/audit_reports/details.cfm?rep_id=3159 – dairy 2014 : http://ec.europa.eu/food/fvo/audit_reports/details.cfm?rep_id=3377

[642] See WHO Global Health Observatory data repository: http://apps.who.int/gho/data/node.main.688

[643] See in World Bank data repository: http://data.worldbank.org/indicator/SP.DYN.LE00.IN

[644] See World Bank data repository, GNI per capita (PPP in current US$): Azerbaijan 2000 = 3,340 – 2014 = 16,910. By comparison, Georgia 2000 = 2,690 – 2014 = 7,510 and Armenia 2000 = 2,380 – 2014 = 8,450. The impact of war (internal secessions, external war with Russia) and of Russian embargo (on key export products) is highly visible in Georgia and has severely curtailed growth. Data available at: http://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD

[645] In 2007-2008 for instance the OIE warned about potential dramatic impact from disbanding the previous system of veterinary inspections. The dire warnings did not as yet materialize (which does not mean nothing could happen, of course).

A last example can be taken from the tax inspection field. As part of its 2007 *Business Environment in Ukraine* report, the IFC of the World Bank Group conducted several calculations to look at comparative "tax yields" in jurisdictions with varied tax inspection levels, both inside and outside of Ukraine. The international comparison (p. 56) showed Latvia as having a far better "yield" (compared to tax potential) by employee with far less inspections than Ukraine. The internal comparison (focusing on SMEs) showed that, between different regions of Ukraine, there was no link whatsoever between the tax revenue per SME and the percentage of SMEs inspected. Additional research showed that most tax inspections in Ukraine brought negligible revenue, while a very small percentage (10% at most) brought roughly 90% of the additional tax assessments and penalties.

As we have acknowledged from the start, none of these data points is enough to fully prove the case, data reliability, attribution and other issues being limitations that cannot be overcome, at least at this point. Taken all together, however, they build a strong picture of how the volume of checks, their coverage and frequency, are essentially uncorrelated to public welfare outcomes – be they safety, health or tax revenue. Equating "more inspections" or "more stringent enforcement" with "higher effectiveness", as is still too often done, is simply not supported by evidence. While one could argue whether the findings from Former Soviet countries can be transposed to an EU or OECD context (considering in particular the different situation in terms of petty corruption – though it is far from unknown in EU or OECD too), these findings should be a warning to those who think that "stronger enforcement" is a priority in developing countries and emerging market, in front of problems such as environmental pollution, or health and safety issues. Indeed, sometimes at least, *more effective* enforcement may be direly needed. This does not mean, however, that "more" or "stronger" enforcement will prove to be effective.

c. Short overview of a few EU countries: Lithuania, France, Italy

As indicated in the introduction, the scope of this research did not allow us to rely exclusively on case studies, nor to undertake a comprehensive review of each and every case that would be possibly relevant. Our aim was not to demonstrate with full certainty a causal relationship (one way or another) but to challenge established assumptions and to check whether, based on available data and findings, some evidence in support of risk-based inspection practices and "smart inspections" could be found. The above examples appear to rather strongly support our hypotheses: there clearly is no positive correlation between "more inspections" and "increased safety", and Britain (using risk-based targeting, risk-proportional enforcement, and putting much emphasis on guidance and support) achieves significantly better results (and/or similar results with far less costs) than countries relying on more "traditional" approaches to inspections and enforcement, such as Germany or France.

Similar analytical work would need to conducted in other regulatory areas to confirm these preliminary findings, but it is rarely easy to find reliable effectiveness data (as we have discussed above at some length), and consolidated statistics on inspections are also hard to come by. Taking a short look at some of available data points from other countries can, however, be done in order to this review of evidence from the practice.

i. Lithuania – OSH, reforms and challenges

In Lithuania, as we have seen, post-2010 reforms led to a significant decrease in the overall frequency and duration of inspections. They also led to important changes in enforcement practices. Because Eurostat standardised data on fatal occupational accident rates is, as explained above, one of the most reliable

indicators to compare effectiveness, it is worth considering briefly labour inspections in Lithuania and their evolutions.

First, the country's performance in terms of occupational safety appears overall poor. The standardised incidence rate for fatal accidents[646] was 6.62 in 2008 (4.56 excluding traffic-related accidents) and only slowly and unsteadily crept down (5.06/3.79 in 2013). By comparison, the EU-28 average is 2.22 in 2013 (1.3 excluding traffic-related accidents), France (which we saw has practices in sharp contrast to Britain, and worse-than-average performance) achieved 3.71 (2.94) in 2013, and Germany 1.29 (0.81). As we have seen, year-on-year variations should be treated with caution (because of the small number of observations, which can result in sharp variations in the incidence rate), but the 6-year trend is clear: poor performance, with slow improvements[647].

Second, inspections pre-reform appeared to be at a relatively high level in terms of frequency, and to be rather "heavy handed" in terms of enforcement, at least compared to the British and German cases studied above. The Lithuanian State Labour Inspectorate's (SLI) annual reports[648] give data on the number of inspections that is not easy to compare, because they include in the "total number of businesses" under supervision all farms, that amount to more than 50% of the total (but with the breakdown only available for most recent years). Unfortunately, the SLI does not indicate how many inspections were conducted in farms versus non-farms. 2012 data, for instance, would suggest a rate of approximately 5% of businesses inspected (if taking the entire population, farm included) but as high as 11% if assuming that nearly only non-farms were inspected – and even more if assuming that mostly *private* businesses (rather than "all economic entities")[649] were visited. Comparing with data from business surveys conducted by the Ministry of Economy (see OECD 2015 b, p. 123) show that in 2012, close to 15% of surveyed businesses reported receiving at least one labour inspection. While comparisons are complex due to different economic structures and uncertain quality of the data, such a coverage would be more comparable to Germany's than to Britain's, particularly considering that the 2012 rate already reflected a strong decrease (from 16,000 inspections in 2008).

In terms of enforcement, in 2008 the Lithuanian SLI issued 10,980 improvement notices, far more than the British HSE, in spite of covering a considerably smaller economy, and nearly as much as in Germany (12,693 for 2008, cf. Tilindyte 2012 p. 192). In that same year, it suspended operations over 2,400 times (against approximately 3,000 in Britain, again for a vastly larger economy), and issued or proposed 2,500 administrative penalties (twice more than in Germany). These numbers suggest that, in addition to a rather heavy coverage by inspections (possibly reflecting a lack of risk-focus), the enforcement approach was heavy on sanctions and rather "confrontational". Data from 2014 suggests there has been a significant change, with only 1,723 notices (down from still 5,192 in 2012, meaning the decrease has accelerated in recent years), only 43 suspensions of work activity, and 511 administrative fines. The collapse in the number of suspensions was particularly marked in 2012 (only 9, down from 230 in 2011) – and it is likely that the uptick in fatalities in that same year led to a partial reassessment (with approximately 40 suspensions per year since then).

This data, combined with the review of practices done by the OECD (2015, pp. 127-130), suggests that the SLI has strongly changed its practices since 2011, in reaction to reform efforts by the Government, to its own assessment of methods and results, and to the economic crisis (which to a large extent explains the collapse in suspensions of operations[650]). The decrease in inspections coverage is less clear (OECD 2015 b, p. 123:

---

[646] Eurostat data available at: http://ec.europa.eu/eurostat/web/health/health-safety-work/data/database

[647] Lithuania has the EU's second-worst performance on fatal accident rates (including or excluding traffic accidents), just after Romania. Neighbouring Latvia has rates that are nearly as high, suggesting long-term trends linked to the post-Soviet context.

[648] Available for 2008-2014 at: http://www.vdi.lt/English/VDI_English.aspx

[649] See detailed enterprise statistics on the Statistics Lithuania website, at: http://osp.stat.gov.lt/en/temines-lenteles51

[650] Both because of the slowdown in the high-risk construction sector, and because of a deliberate policy by the SLI to minimize suspensions – direct interviews with senior SLI officials, december 2014.

decrease in 2013, but uptick in 2014), but seems nonetheless real over the last 6 to 7 years. At the same time as the SLI thus reduced its inspections coverage (at least somewhat) and turned away from heavy sanctioning to embrace a far more "compliance supporting" enforcement approach, key effectiveness indicators showed overall a slight improvement. Even the 2012 "surge" in fatalities just brought their level back to 2008. Thus, it seems that at the very least it is possible to conclude that the move towards a far more risk-proportionate and more compliance-supporting approach, while it clearly meant far less burden on private businesses, did not have any negative results on effectiveness – and probably had some limited positive impact. It is still too early to say whether time will strengthen this positive trend, but it is at least clear that the heavy-handed approach used for many years did not have any positive impact in terms of safety, and that a change was clearly in order.

ii. Labour inspections in France: conflicts, data issues, and disappointing results

Many "old Member States" of the EU (i.e. those who joined before successive "eastern enlargements", up to 1995) have disappointing OSH results, i.e. data that is significantly worse than the EU-28 average. This is the case for instance of Luxembourg, Belgium or Spain if excluding traffic-related accidents – and of Ireland, Italy or Austria if including them[651]. We elected to look briefly at only one of these, namely France, because labour inspections are a highly contentious issue there, with important political forces considering that any decrease in frequency of inspections or any laxity in enforcement would have dramatic consequences for workers (and other forces seeing labour inspections as a major impediment for business development). While the fatal accidents data is clear, however, it is not the case of inspections data. The Ministry of Labour publishes yearly reports to the International Labour Organization that include relatively detailed statistics on labour inspections and enforcement[652], these statistics are based on reporting in a unified information system by labour inspectors, and the annual reports repeatedly point out the variations in reporting rates and accuracy of reporting[653], meaning that the data should be considered as more indicative than authoritative. In addition, as we have pointed out above, other institutions inspecting OSH issues are not accounted for in these reports, which mean they seriously under-estimate the overall inspection and enforcement activities in France on this issue.

Considering the uncertainties surrounding the data, it is difficult to make definitive assertions, but there seems to be no real trend in terms of frequency of inspections, or of enforcement measures, over the period 2000-2013. According to the annual reports for the years 2000, 2003, 2008 and 2013, the key indicators were as follows:

	2000	2003	2007	2011	2012	2013

[651] Austria has a particularly high fatal accidents incidence when including traffic-related accidents, averaging more than 5 over 2008-2013 – since its performance excluding traffic accidents is just barely worse than average, this suggests a specific and very acute transport-related problem.

[652] See links to reports since 2000 on the Ministry's website, available at: http://travail-emploi.gouv.fr/ministere/documentation-et-publications-officielles/rapports/article/l-inspection-du-travail-en-france-en-2013

[653] See in particular the 2000 report: *l'inspection du Travail en France en 2000 - les chiffres clés. Rapport au Bureau International du Travail*, Ministère de l'emploi et de la solidarité, Paris, 2000. This report indicates that all statistics on inspections and other interventions published in the 1990s are essentially worthless because of erratic reporting rates and low reporting accuracy (p. 190).

Number of visits to businesses[654]	295,930[655]	295,899[656]	239,542[657]	356,200	265,300	294,000
Official letters (warnings)	N/A	N/A	161,114	226,300	163,000	183,500
Improvement notices (*Mises en demeure)*	9,621	7,921	5,017	6,573	5,515	5,375
Prosecutions	N/A	N/A	N/A	8,345	7,624	6,374

While there seems to be some decrease in the most stringent form of improvement notices (*Mises en Demeure*), there appears to be no noticeable trend in any of the other indicators (some of the data is very differently reported in different years, explaining the repeated "not available" mentions). The quality of reporting data seems to be a real issue, with software problems compounded by inspectors' refusal to use the new procedures. The 2003 report (pp. 220-221) outlines a long-term decrease in the total number of interventions, but this appears based on incomplete data, and partly reversed in recent years. The high number of warnings and prosecutions (compared e.g. to Britain or Germany) reflects the overall climate surrounding labour inspections in France, characterized by distrust on both sides (inspectors and employers).

During the same period, no improvement trend is visible in our key safety indicator, i.e. fatal accidents incidence, as the table below demonstrates[658]. France has the 9th worst average over 2008-2013 if including traffic-related accidents, the 7th worst if excluding them. It seems to be sliding in relation with the performance of other "old Member States", since over 1998-2007 had the 11th best performance – which meant 4 of the EU-15 had worse performance. Since 2008, only Portugal (and Austria, if including traffic-related accidents) has worse performance among the EU-15 group.

Eurostat Data	**Fatal occupational accidents, France**								
	1998	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**Average (98-13)**
Pre-2007 ***(excludes traffic accidents)***	4	3.2	2.6	2.8	2.7	2	3.4	2.2	**2.97**
			2008	**2009**	**2010**	**2011**	**2012**	**2013**	**Average (08-13)**
2008 onwards ***(excludes traffic accidents)***			0.5	2.07	2.59	4.99	2.64	2.94	**2.62**
2008 onwards ***(includes traffic accidents)***			1.84	2.9	2.91	8.11	3.51	3.71	**3.83**

654 In France, there are many types, among which "control visits" (inspections *stricto sensu*) are only the most frequent, around 60% according to the most recent annual reports. For greater simplicity, and because the proportions appear more-or-less stable, we use here the total number of visits, whichever their legal nature and cause.

655 Extrapolated from the report's figure of 216,029, based on the note that this represents approximately 73% of the real number of interventions, cf. p. 162 of the 2000 Report to the ILO. Extrapolation for improvement notices is based on indication (p. 163) that only 55.7% of agents reported on this indicator.

656 Extrapolated from the report's figure of 253,386, based on the note that this represents approximately 85.7% of the real number of interventions, cf. p. 219 of the 2003 Report to the ILO. Same extrapolation done for all data in the 2003 column.

657 Extrapolated from the report's figure of 215,588, based on the note that this represents approximately 90% of the real number of interventions, cf. p. 155 of the 2008 Report to the ILO. Same extrapolation done for all data in the 2007 column.

658 As for data presented above on Britain and Germany, there is a break in time series after 2007, meaning that data up to and after 2007 is not directly comparable.

What these findings suggest is that OSH inspections in France seem to suffer from sustained problems: poor data (making it difficult to track evolutions), lack of questioning of prevailing practices and their effects even confronted with repeated poor performance, high reliance on formal enforcement measures but without evidence of positive impact on compliance. Data suggests that there is neither a trend in reducing the inspections incidence or the amount of enforcement measures, nor a positive evolution in safety. This suggests that the reliance on a "traditional", confrontational and non-risk-proportional approach does not seem to yield positive results.

iii. The inspection system in Italy – structures, crises, attempts at reforms

The survey conducted in Italy in 2012 on businesses' experiences with inspections offers a level of insights on inspections coverage and patterns that is unfortunately available in very few countries[659]. There are, however, some limitations in the uses that can be made of it, at least without considerable further research. Since no subsequent survey was conducted (or at least published), it is not possible to check for any changes in the most recent years. As most other EU countries have no comparable data, it is not possible to directly compare levels of inspections or patterns. And, because of all the limitations in effectiveness data discussed above, it is very challenging to easily assess the effectiveness of these inspections. That said, there are a number of interesting points in the data, which at least allow to confirm the relevance of some of the questions raised in this research.

In terms of coverage, in spite of inspections being conducted by a large number of agencies, in a variety of fields, the bulk of the control visits (and also documentary controls) are done by a small sub-set of them, focusing on a narrow set of domains: food safety and public health (*Aziende Sanitarie Locale* – ASL, Local Public Health Establishments, controlled the highest number of businesses), fiscal issues (the *Guardia di Finanza,* i.e. tax and customs police, was second, and the tax agencies fourth – cumulated, they controlled more businesses than even the ASLs), OSH and labour law (Labour Inspectorate and INPS, *Istituto Nazionale Previdenza Sociale* – National Institute for Social Prevention plus INAIL, *Istituto Nazionale per l'Assicurazione contro gli Infortuni sul Lavoro* – National Institute for Insurance against Labour Accidents – again, the combination of Labour Inspectorate, INPS and INAIL totalled more controls than the ASLs). A fourth group was made up by controls of the various police forces (excluding the *Guardia di Finanza,* GdF), which can cover a variety of issues (food and hygiene, environment, "nuisances" and public order etc.). In total, there were (in 2011, for businesses with 5 to 250 employees) more than 115,000 businesses controlled tax and duties, around 100,000 for labour and OSH inspections, over 80,000 by ASLs, close to 50,000 by police forces (excluding GdF) – but only around 10,000 by the Regional Environment Agencies (*Agenzie Regionale per la Protezione Ambientale*, ARPA), around 17,000 by the Fire Service, and less than 10,000 by the Forest Corps.

Many of these inspections in fact covered repeatedly the *same* businesses, as we have already pointed out earlier. Data shows a considerable amount of overlap between ASLs, Labour Inspectorate, GdF, tax service and INPS. Thus, a given business, if it was inspected, was likely to receive repeated visits on very closely related issues, or even on the same subject, from different (unrelated and uncoordinated) agencies. While some

[659] The survey results were presented in several public events in April and May 2012, as well as later in the year, but are not available in a published form. The below is based on data distributed during these presentations, which the author attended. The survey was conducted by the Italian State Statistics Agency ISTAT on behalf of the Office for Administrative Simplification in the Department for Public Administration under the Presidency of the Council of Ministers
(see their website at: http://www.funzionepubblica.gov.it/uffici/ufficio-la-semplificazione-e-la-sburocratizzazione)

reforms were initiated to try and address this situation (e.g. by taking away the inspection function of the INPS), it is unlikely that this situation has altogether changed since 2012.

As a result, it appears that inspection resources are in many instances fragmented and not used in the most efficient way (with repeated controls on the same topics by agencies that do not share planning or results), a concentration on some issues (e.g. labour/OSH) that produces mixed outcomes (as we have seen above, Italian performance in OSH is below EU-28 average at least when including traffic-related accidents – when excluding them, it is below EU-28 average, but still above EU-15 average[660]), and possibly a lack of resources for some important risk areas. This is particularly the case of environment, which we discuss briefly below. To conclude on the data, however, it is also important to note that the survey also included questions on the number of hours required to deal with inspections by each authority. These showed without surprise that tax and duties controls took the most time (1 to 2 days each), but the average time for labour-related and health-related inspections (around 5 hours each) was also not insignificant, particularly considering that the survey focused on SMEs.

Several major scandals affecting environmental protection, public health and occupational safety and health have broken out in recent years in Italy – and both are still ongoing issues. The first is the waste-management crisis in Southern Italy (leading to very serious public health concerns), and the associated food safety issue when it was found that illegal waste disposal had led to dioxin-contamination in *mozzarella di bufala.*[661] The second is the pollution scandal linked to the Ilva Taranto steelworks (long the largest in Europe), which exceeded applicable norms for decades, with major effects on public health in the area (both workers' and general population's) and the environment more broadly.[662] At first glance, these scandals could point towards some serious gaps in environmental protection inspections and enforcement – however it is not clear how much they are linked to what one could call "regular" inspections and enforcement, or whether they rather reflect very specific political and criminal contexts, on which regulatory agencies have very little influence (if any). In Campania and other Southern Italian regions, indeed in principle regular inspections could have helped to spot problems early on (particularly illegal waste dumps), but it is unclear how much inspectors could have done in a context where such illegal waste operations were managed by organized crime. As to the lack of adequate investment in proper, legal waste management, this went back to political decisions (and, again, criminal influence) – not issues on which any "business inspections" could have helped. The Ilva Taranto case is different[663]: environmental inspectors did their work, found out about the (major) violations, notified the need for improvement, attempted to withdraw the environmental permit allowing the factory to continue operating – but the political connections and wealth of the Riva family (which owned the plant) and the huge social importance of the plant (the main employer, by far, for Taranto and its region) meant that there was constant political backing (from otherwise opposite political camps) to adopt special legislation allowing it to continue operating. Only once the case was taken over by criminal prosecutors did it become possible, as part of the criminal case, to suspend operations. Thus, the problem was not the environmental inspectors failing at the task (they did not), but politicians overriding them (for a variety of reasons). Both of these cases show

---

[660] Though Eurostat data on fatal accident rate suggests a trend for improvement – the 1998-2007 series (excluding traffic accidents)saw Italy's rate decline from 5 to 2.5, and the 2008-2013 series a decline from 4.5 to 3.06 (including traffic accidents) and from 1.89 to 1.24 (excluding them). It may thus be that some results are seen from sustained efforts, but the overall performance remains worse than EU-15 average (though better than results in France). The Ilva Taranto disaster (see below) also shows that long-term occupational health risks may be a serious issue at least in some parts of Italy.

[661] See Pasotti (2010) on the waste management crisis, and Borrello, Brambilla, Candela *et al.* (2008) on the mozzarella contamination scandal.

[662] See Pascucci (2013)

[663] There exist many summaries of the case, which has developed over a couple decades – see e.g. the Wikipedia article in Italian, which is regularly updated and has a number of links (https://it.wikipedia.org/wiki/Ilva#Taranto) or articles in *La Repubblica* such as http://www.repubblica.it/ambiente/2015/03/03/news/good_morning_diossina_il_libro_di_angelo_bonelli_sul_caso_taranto-108646777/ or http://temi.repubblica.it/micromega-online/ilva-uno-scandalo-di-incompetenza-e-malapolitica/?refresh_ce

how important it is not only to distinguish between "regulatory enforcement" and "criminal enforcement", but also the importance of having *effective links* between them – as there are cases which cannot adequately be addressed from a regulatory compliance perspective.[664]

Considering available data on key performance indicators e.g. in food safety[665] or environmental protection[666] does not suggest that Italy has very serious structural problems of regulatory compliance "across the board". Still, performance in food safety (at least from EFSA/ECDPC reports) is not outstanding, and neither is environmental performance. Occupational safety performance is, as we have seen, below par. Considering that this happens against a background of inspections that are overall rather frequent, and particularly "concentrated" on a limited number of businesses (many repeat inspections), with significant institutional overlaps, there seems to be a real case for improving risk analysis, targeting, compliance promotion methods, and overall coordination of inspection activities. All these areas belong to the reform work that the Government has undertaken since 2012, but it is still too early to see whether the situation has changed significantly.

4.2. Data challenges – inherent limitations in considering factual evidence

As we have briefly discussed earlier in this research, the question of whether regulations pose a significant burden on economic growth and competitiveness cannot be fully responded to based on available research. Investigating whether regulations deliver their expected benefits in terms of public welfare is likewise complex, and existing research gives conflicting answers (which often reflect different regulatory approaches and goals, but may also correspond to limitations in research design). The focus of our research is markedly more modest and limited: attempting to find out whether different approaches to inspections and enforcement appear to have different effects in terms of public welfare, while also considering the level of administrative burden they create (which is an admittedly very imperfect proxy for the impact on business growth, but is an indicator that can be more-or-less easily available), and the degree to which they include compliance-supporting activities or not. All through the different cases exposed above, we have relied on

---

[664] Note that such effective interaction can be through integration of prosecution in the regulatory agency itself (British HSE) or by coordination between distinct institutions.

[665] See for instance the *European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2014*, European Food Safety Authority and European Centre for Disease Prevention and Control, available at: http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/4329.pdf (and see generally EFSA monitoring and other reports at: http://www.efsa.europa.eu/en/topics/topic/monitoringandanalysisoffood-bornediseases). While EFSA and ECDPC warn against limitations in comparability, and there are many indicators for which reporting bias is an issue, data from monitoring campaigns done according to specific EU regulations (where comparability is maximal) do not suggest any systemic weakness in Italy – e.g. in terms of *Salmonella* control, it overperforms in some areas (e.g. contamination in breeding flocks, p. 44, is below average), underperforms in other (e.g. contamination in laying eggs flocks, p. 47, is somewhat over average). The situation is somewhat worse in broiler flocks before slaugher (pp. 48-49). Overall, Italy is rarely among 'best performers' (though it is in some indicators), but the situation does not appear to be cause for concern either.

[666] Italian compliance levels in key areas such as EU regulations on water and air do not seem out of line with other Western European Member States. For instance, the most recent report on water quality is Synthesis Report on the Quality of Drinking Water in the EU examining the Member States' reports for the period 2008-2010 under Directive 98/83/EC (COM(2014) 363), available at: http://ec.europa.eu/environment/water/water-drink/pdf/report2014/1_EN_ACT_part1_v3.pdf shows Italy having no problems in microbiology, and some (isolated) issues in chemical contamination, but not more than Germany or Spain (pp. 4 and 10). There is no equivalent report for air quality, but the closest comparison we found is the In depth analysis of the NEC national programmes - Final Report, prepared by ENTEC (2005), available at: http://ec.europa.eu/environment/archives/air/pollutants/pdf/final_report.pdf. The Executive Summary (page iii) states that "only four Member States are currently projected to comply with all of their NEC targets by 2010 without the need for further actions" – suggesting a far lower level of compliance with the targets. Specifically for Italy (pp. 87-88), the report highlights some limitations on public action, but again not different from what is found in many Member States, and reflecting broad public policy issues rather than specific shortcomings in inspections and enforcement.

aggregate, country-level, average data, and not attempted to do any statistical analysis of possible correlations – be it across countries (using country-level data) or within a country (using e.g. firm-level data).

There are several reasons for this. First, the purpose of this research is rather to look at the issue in as broad a way as possible, considering a large number of countries and regulatory areas, and combining a review of existing research and explanatory models with evidence from the practice. The time and efforts required to attempt a statistical analysis of the data would be substantial, and would have required to curtail other parts of the research. Second, and more fundamentally, there are reasons to doubt that such work would produce conclusive results. These doubts are grounded both in the fact that existing statistical research on the effectiveness of inspections has yielded rather conflicting research, and on the many limitations and shortcomings of the datasets that could potentially be exploited. Before concluding on this chapter considering examples from the practice, we will briefly present some of the examples that led us to avoid engaging in statistical analysis of data (in spite of the benefits it could potentially provide if conclusive results were within reach).

a. Limitations and contradictions in studies attempting to investigate effectiveness

As a preliminary remark, it is essential to note that existing data-based research purporting to investigate the effectiveness of inspections was formulated as looking into the effectiveness of "regulation", with inspections seen the primary means of translating this regulation into practice, but not considered in their specificities. Thus, this research did not specifically look at the methods used to target these inspections (and whether some may be more effective than others), or the approach used during inspections (and whether they may have different impact), but only at whether there was a statistically significant effect of having had an inspection, versus not having had any. Given that the topic of our investigation was to see whether there was a *differential* impact between different methods and approaches for targeting and conducting inspections, the findings of these studies are anyway interesting but not directly conclusive. Still, they are interesting as an illustration of the pitfalls existing when attempting to base such research on strict statistical analysis.

The studies we will now consider focus on occupational safety and health, a regulatory area that has been subject to a number of studies, primarily in a US context, largely because of its high level of "political salience", with political parties sharply divided about its costs to the economy, positive impact on welfare, and overall policy choices in this regard. Because many other countries have less of a political conflict on this topic[667], there have been less such studies elsewhere. This means that in many other countries researchers have rather assumed that having occupational safety and health regulations was in and of itself likely to have positive effects on workers' health and safety, and research has thus focused on the approach and "style" of regulation and enforcement, but not on whether having regulations (as opposed to having none) had an impact, and which ones[668].

Interestingly, two significant studies on this issue have very similar approaches – but end up with findings that *appear* contradictory at first (we will see that "deconstructing" the findings allows to understand, if not solve, this contradiction) – even though the second survey references the first. We will start by summarizing briefly

[667] see e.g. Clark 1999 on the difference between US and Australia in matters of occupational safety and health regulations – but note that OSH and labour regulations are very "political" and "confrontational" in France and Italy, for instance (as illustrated in Italy by labour inspections being excluded from the scope of the inspections reform that started in 2012 – in France, yearly reports to the ILO regularly discuss physical conflicts involving labour inspections, and the need to request police protection, which is a good illustration of the level of conflict).

[668] This is possibly also because the Occupational Safety and Health Act of 1970 provided a convenient cut-off date for comparative research – before that date, researchers could consider that regulation was relatively minimal, federal inspections nearly absent, and overall the level of regulatory intervention was very low.

the two studies and their findings, before discussing their assumptions and conclusions, and seeing if we can draw some lessons from the apparent contradiction(s).

i. Inspections: effective, but not for their stated purpose?

In a first paper (Bartel and Thomas, 1985), the authors used official data from the Occupational Safety and Health Agency (OSHA) database, on inspections, (non-)compliance findings (and enforcement), and combined it with data obtained from the Bureau of Labor Statistics and Census Bureau, covering workers' injuries rate, firm size etc.. The data used covered the 22 states where *only* OSHA was enforcing the 1970 Occupational Health and Safety Act (and not the 28 others where state-level enforcement was also involved). The study intended to test two conflicting hypotheses on why previous studies had "failed to find any statistically significant impact on national injury rates due to the Occupational Safety and Health Administration" (p. 1) – the first is the "noncompliance hypothesis" ("because of limited statutory and budgetary authority from Congress, OSHA is unable to compel industrial compliance with its own standards") and the second the "inefficacy hypothesis" ("Since OSHA standards address only part of the problem, these standards can have at best minimal effect[669]") (pp. 1-2). The authors aim to explain what they consider a paradox – that in spite of apparent ineffectiveness (since there is no impact on national injury rates), and in spite of "enormous financial burdens on industry" , "OSHA safety regulations and their enforcement were continuously supported and funded by Congress throughout the decade 1970-80 despite significant Congressional controversy" (p. 2).

Among the study's major findings are that "the negative and significant coefficient on the inspection probability (…) indicating the responsiveness of firms' compliance decisions to OSHA's enforcement efforts" and that "noncompliance is also strongly affected by increases in the penalty structure (…). Indeed, increases in the penalty structure are a more efficacious means of achieving greater compliance than increases in inspection rates. Hence we have quite strong evidence that the noncompliance hypothesis is false" (pp. 20-22). However, "the result of a doubling of the inspection rate is only a 2.5 percent reduction in the lost workday rate because of the weak relationship between compliance and safety" (p. 22). The authors further note that "large firms (…) clearly choose lower violation rates because of lower marginal costs of compliance. These findings demonstrate the presence of significant economies of scale in compliance for large firms and, therefore, the opportunity for redistributions of wealth from small to large firms through OSHA enforcement" (p. 22). Finally, they note that "industries with higher injury rates (holding constant compliance levels) and industries with higher profit rates are inspected more frequently" and that "unionization has a negative and significant coefficient; this implies that unionized firms use OSHA as a tool for imposing costs on nonunionized firms. In addition, industries with larger average firm sizes have lower inspection rates, although this enforcement asymmetry disappears by 1978[670]".

The authors conclude that "our study has found only weak linkages between noncompliance and workplace accidents, indicating that the inefficacy hypothesis is largely correct, although the statement that OSHA standards achieve no reductions at all in injuries is probably invalid. In contrast, there are significant effects of OSHA enforcement on industry violation rates, indicating that the noncompliance hypothesis is false" (p. 25). In other words, they find that inspections and sanctions are effective at increasing compliance, but that the rules are inadequate and thus compliance ends up having little positive impact on safety. Their second conclusion is that "indirect effects of OSHA regulations exist, are significant in magnitude, and may well

[669] "It is important to recognize that OSHA standards are not performance requirements that specify some maximum accident rate for each firm, but rather are design requirements for the workplace itself. Most OSHA standards are in fact capital equipment standards dictating, often in great detail, physical characteristics of plant and equipment." (p. 4)

[670] Possibly because of greater political attention given to complaints by small businesses, referred to in the study.

dominate any direct effects (certainly direct benefits). The apparent beneficiaries of these indirect transfers of wealth are unionized and large firms, who would reasonably provide political support for the agency, so long as OSHA has some cost impact-and so long as this impact is asymmetrically distributed against nonunion and small firms" (p. 25). The overall conclusion of the study is thus to validate a "regulatory capture" vision of OSHA standards and enforcement strategy, as well as a "deterrence" vision of inspections (but with a lack of effectiveness to improve safety, the stated objective of the rules and of the institution).

ii. Inspections: effective, and with "lagged" effects?

In a second paper (Scholz and Gray 1990), the authors built upon past research (including Bartel and Thomas, which we just summarized), but were able to use a different dataset "merging OSHA enforcement records and Bureau of Labor Statistics (BLS) injury data for 6842 manufacturing plants", which "provides richer information than has been available to other studies" (p. 284). The data covers 1979-1985, i.e. a different period than the previous study. It combines data on inspections and enforcement actions, characteristics of the plants inspected (including size, but also some qualitative data on the workforce), injuries (not aggregated rates, but at the plant level). The study was predicated generally on the same deterrence-based compliance model (as originally formulated in Becker 1968), but incorporating major elements from the "behavioral theory of the firm" (p. 283), as well as findings from research on "decision making under risk and uncertainty" (coming from Kahneman, Slovic and Tversky 1982 in particular) (p. 285). Thus, the model being used is somewhat more sophisticated (in that it incorporates more findings on behaviours and decision making), but still founded on deterrence theory. The authors investigate four hypotheses: that firms respond to accidents (by attempting to correct the safety risks they have revealed), that firms respond to perceived increases in "enforcement risk" (OSHA enforcement activity) but "over several years", that OSHA imposing penalties against firms has an additional, "specific deterrence" effect – and that firms react independently to the two "dimensions of expected penalty (probability and amount)" and "respond more to changes in probability".

A key difference with other studies (and one that the authors think explain a substantial part of the difference in findings) is that the sample is not representative of the general enterprise population but, rather, corresponds to types of plants which are an area of focus for OSHA. This is due to the BLS over-sampling large plants. As a result, the sample had firms with on average nearly 10 times more workers than the general enterprise population, and far more regularly inspected ("27% of them inspected in 1979, compared with 8% for all manufacturing plants", p. 288). While not representative, this lent the sample "more analytic power" (*ibid.*) to try and investigate responses to enforcement activities.

In their conclusions, the authors note that "the number of lost workdays and (…) injuries decrease significantly after increases in general enforcement and after specific contacts with enforcement agencies" (p. 302), i.e. there is both a general effect of OSHA enforcement existing (and increasing), and a specific effect of OSHA visits (the authors consider this a deterrence effect but one could argue that it may be a broader effect, and not only deterrence). They also note "relatively long lags between enforcement changes and changes in injury risks" (p. 302), and that "our estimates suggest that the effect [of changes in enforcement] continues into later years" (p. 295). In addition, "the results confirm that changes in probability and penalty are not symmetrical (…) increase in inspections reduced injuries and lost workdays more than a comparable increase in penalty" (p. 297). However, "enforcement effects are relatively modest, as other studies have found; a 10% increase in enforcement would reduce injuries by around 1% for the large, frequently inspected firms represented in our sample" (p. 302). In short, while inspection effects are small, they are clearly present, and produce effects over a couple of years (which can be both because the behaviour effects last a couple of years before reverting to mean, and because the positive effects of capital investments in safety take some time to produce results).

iii. Making sense of the findings: "how" rather than "whether"?

At first glance, the difference in findings may partly be explained by the sample composition and more detailed data: Scholz and Gray had more detailed data (with injuries etc. *per plant*), and in any case the effects are small, thus looking at a sub-sample of larger, more frequently inspected businesses, magnifies results that may otherwise be too small to be significant. Revisiting these two studies from the perspective of practical experience in inspections across many countries allows us, however, to challenge a number of the underlying assumptions for the two studies, and to suggest possible alternative interpretations of their findings.

First, both studies rely fundamentally on the same model of compliance – though a slightly modified (or "enhanced") version in the case of Scholz and Gray. The fundamental driver is seen as deterrence, with compliance entirely (or mostly) determined by rational calculations of costs and benefits. For Barthel and Thomas, "firms will elect to violate OSHA standards whenever such noncompliance is profit maximizing. Even apart from OSHA enforcement efforts, the level of noncompliance by a firm will have several distinct effects on profits" (p. 5) – probability and cost of enforcement actions intervene by modifying the cost-benefit calculation. In their study, Scholz and Gray use the same model but with additional considerations for "risk-induced" behavioural responses (reaction to accidents etc.), "based on observations of business decision making processes, in particular observing that firms' behavior deviates systematically from optimal performance (which would simultaneously maximize expected profit over all possible behaviors) because of limitations on the firm's decision-making ability" (p. 286). There is no consideration that firms' behaviours (and the behaviours of individuals who work within the firm) may be driven by a variety of other factors (social, cultural, psychological etc.), values, and thus may not be strictly determined by a combination This is all the more striking considering that, as the authors themselves write, "most empirical studies have investigated the deterrence hypothesis" but "the results of these studies have not consistently demonstrated the linkage [between deterrence and compliance] (…), although the insignificance of effects is sometimes interpreted as a sign of ineffective or inadequate enforcement rather than of a weak theory" (pp. 283-284). The inconsistency of findings, as well as the modest magnitude of deterrence effects observed even in Scholz and Gray's study (in spite of its sample "bias" towards more-heavily inspected firms) would, rather, appear to us to strongly support the view that deterrence is *not* the only or even the main compliance driver (at least in most cases), and that the modesty of the observed effects simply reflects this. Putting too much emphasis on deterrence, regardless of the sophisticated economic models developed for these studies, flies in the face of daily evidence. Every day, most individuals will comply with rules and norms for which the probability of detection, were they to violate them, is vanishingly small – and thus the deterrence effect very low[671]. A more sophisticated model of compliance is clearly necessary.

Second, while the authors have attempted to control for a number of factors, the way they have done it is not entirely convincing, which has to do with their overly schematic compliance model, with insufficient consideration of the specifics of the phenomena studied, and with data limitations. Because of the deterrence model, they did not look at the potential influence of factors such as information about regulations, type and quality of interactions with inspectors. Because of data limitations (or at least so we have to assume), they did not look at the question of costs of compliance, and whether there appears to be significant differences in compliance between firms which would have substantially different costs (in fact, both studies repeatedly refer to the question of costs, but always with assumptions and never with data). Even on factors for which they *do* attempt to control, one can but notice that they do so with some lack of attention for the specifics of

[671] Quoting Tyler (2003): "In most actual situations, the objective risk of being caught and punished is quite low. For example, according to an analysis of crime and arrest rates, the objective risk of being caught, convicted, and imprisoned for rape is about 12 percent; for robbery 4 percent; and for assault, burglary, larceny, and motor vehicle theft 1 percent". Even if a number of people over-estimate the probabilities, the fact that the vast majority of us do *not* steal cars is sufficient evidence that a large part of compliance cannot be explained by deterrence effects only.

the issue. A perfect example is the question of reporting of accidents and injuries. This is a fundamental variable for both studies, but barely gets any discussion. Bartel and Thomas write that "an analysis of worker injuries must take account of the role played by the workers' compensation system. The benefit structure varies across states and over time, and previous research by Butler and Worrall has shown that reported injury rates are higher in those locations and those years when benefit formulas are the most liberal" (p. 13) and they therefore opted to correct for this by using the "expected benefit measure for a representative wage earner" in each industry, and a "weighted average of the waiting period for receipt of benefits". Scholz and Gray make no mention of this issue at all. However, we know from many other settings and studies (e.g. Tilindyte 2012 pp. 122-124) that there are many situations of under-reporting of incidents, for a variety of reasons (mostly employers' pressure and/or attempts to reach an informal settlement and avoid any possible liability, sanctions etc.). There may be reasons why these do not apply to the US, but none of the authors even discuss them.

Third, the two studies make a number of somewhat "heroic" assumptions and/or downright *non sequiturs*. A strong example of this is to be found in Bartel and Thomas (p. 14): "Violations of OSHA standards are much like victimless crimes in that they are not automatically reported, but rather must be uncovered and verified by inspections. Not violations per firm (VF) but only registered violations (R) generated by inspections (I) are observable. The variables are related as follows: R = VF I. Hence registered violations per inspection (observed noncompliance) is a proxy for violations per firm (actual noncompliance)." In fact, this makes a number of (hidden) assumptions and there is no logical link from the premises to the conclusion. It assumes that each inspection finds all violations in the establishment, and that violations (and/or inspections) are distributed randomly so that indeed one variable can be used as proxy for the other. There is no reason to assume that this is correct.

Fourth, the fact that the effects of inspections (on compliance, and safety) are found to be small, but appear to increase (or be more significant) when one "focuses" more (e.g. through Scholz and Gray's sample of larger firms), actually suggests that the line of inquiry should be different. The question is not so much "do inspections achieve anything" but "under what circumstances, with which methods, do inspections work better?" If we take a different (or complementary) interpretation of the persistence of OSHA funding in spite of disappointing aggregate results in the 1970s, i.e. that Congress knew that the public wanted more protection, how do we make this protection more effective? Thus, we would argue that the question asked was maybe not the most relevant one – and that rather than asking "is there an impact from inspections", the question should rather have been "are there more efficient and effective ways to reach the desired impact".

In any case, these two studies show the limits of statistics-based investigations in our field. Rather than yielding conclusive and solid findings, two successive studies considering the very same issue result in largely opposite findings, partly reflecting differences in the data, but also to a large extent differences in the methodology and underlying assumptions (different weights, coefficients etc. given to different variables and phenomena). Considering the very considerable resources required for this type of work, the cost/benefit ratio does not appear very favourable.

iv. Broadening the view: US OSHA in international perspective

Let us attempt to draw a couple of lessons from these two studies, their strengths and shortcomings. Overall, they are "on the verge" of making a significant contribution to understanding inspections' impact on safety but, in our view, do not quite reach that point because they adopt a narrow model and neglect a number of aspects of the problem – and because they remain too far from the practice to actually consider how the institution selects targets, how the inspectors conduct their work.

To some extent, this problem is exacerbated by the specific characteristics of occupational safety and health (OSH) regulation in the US. For a number of reasons (in particular the fact that the OSH reform movement was more linked with broad social movements, environmental and consumer activists etc., than with organized labour), in the 1970 OSH Act "cooperation, discretion and flexibility were designed out, while adversariness and strict adherence to rules and procedures were designed in" (Clark 1999, p. 99). In line with this, "OSHA's deterrence philosophy involves a much greater emphasis on citations and penalties. US compliance inspectors are accorded little discretion" (*ibid.,* p. 96). In addition, "the courts are highly involved in the development and interpretation of US health and safety regulation" (*ibid.*) – which means OSHA will be keen to adopt rules that prescribe as much as possible "objectively verifiable", material standards, rather than practice-focused / safety outcomes requirements, to protect itself against judicial review. At the same time, "the operations of OSHA have continued to attract an intense amount of scrutiny, controversy, challenge and criticism in Congress, the courts, the executive branch (…) and the media" (*ibid.*, p. 98). It would be understandable, in this context, that researchers gave less attention to "enforcement style", since it is so obviously constrained – though in fact Scholz himself has written elsewhere on the importance of enforcement methods (Scholz 1994)[672].

Looking at the practice, here, would have meant trying to understand how OSHA was selecting inspection targets, and if some selection methods gave better results, for instance. It could also have meant comparing OSHA's results with that of other agencies where inspectors are given more flexibility and discretion, to investigate whether the benefits of discretion can outweigh the risk of capture. Even with the very rigid regulatory framework that surrounds OSHA, there is clearly scope for improvements in data analysis, targeting, outreach and information, and development of inspectors' skills – all these can have a major influence on results, and looking only at aggregate results in one agency tells us relatively little. In fact, to the extent that Scholz and Gray's results indicate a stronger effect than Bartel and Thomas's, and that this seems at least to a large extent due to their more focused sample, these studies in a way show "by contrast" that a more focused selection approach (e.g. a risk-based one) would produce more effects on safety – but this point is not really seen, as the focus is on validating a set of theoretical hypotheses rather than on understanding how practice works.

A last point of interest considering the US OSHA case would be whether its practices appears to deliver better, or worse results than other, more "responsive" ones. There is no easily available conclusive evidence on this topic, but we can look for some indications. Ayres and Braithwaite (1992) wrote of OSHA's regulatory strategy that it was "poorly conceived", with inspectors that "constantly nip at firms with flea-bite fines", where "petty punitiveness is in the foreground and no big guns are in the background" (p. 49). By contrast, Clark (1999) wrote that "there is evidence to suggest that, in terms of workplace outcomes, Australia's current occupational safety and health performance is, at the very least, *no better* than that of the USA" (p. 102) – but unfortunately she did not indicate any of this evidence.

The only relatively "easy" indicator that is available to compare the performance of OSH systems across countries is the fatality rate, which suffers from less distortions (and under-reporting) than the rate of (non-fatal) injuries – even though one of the studies used here reports that "limited or incomplete information on the death certificate and variation in certifier interpretation of the "injury at work?" item contribute to an estimated under count of occupational injury deaths of between 10% and 30%" (Feyer, Williamson, Stout, et al. 2001, p. 23). With this *caveat* then that one should not exaggerate the precision of rates of fatal occupational accidents, there are at least four published papers or reports comparing these rates across a

[672] See also Kagan (1989, 1994) for another example of US-based research emphasizing the importance of different "enforcement styles".

number of advanced economies (and online data allows to complement older studies by seeing how rates have evolved).

Feyer *et al.* in 2001 compared New Zealand, Australia and the US, and showed the US as having somewhat better results than Australia (and New Zealand performing worst), but "because the United States data collection method likely underestimates the occurrence of work related fatal injuries, the true difference between the United States and the other two countries is probably less" and "much of the difference between countries was accounted for by differences in industry distribution" (the differences within a given industry being far smaller, and sometimes going in the opposite direction) (p. 26). It should be noted that Australia's performance appears to have improved strongly over the decade and more since this paper was published, as the most recent data shows that "The 191 fatalities in 2013 equates to a fatality rate of 1.64 fatalities per 100 000 workers. This is the lowest fatality rate since the series began 11 years ago. The highest fatality rate was recorded in 2004 (2.94)" (Safework Australia 2013, p. vii).

The next study was done by Australia's National Occupational Health and Safety Commission (2004). It decided to exclude the United States "as their data, particularly at industry level, would require significant manipulation before it could be included" (p. 23). The report added that "as the USA's incidence rate (4 deaths per 100,000 employees per year) is considerably higher than the countries selected, it is unlikely its inclusion would alter the findings of this report. This is in contrast to the report by Feyer *et al*. (2001) which concluded using 1989–92 data that the USA performed better than Australia. Since this time Australia's performance has improved substantially whereas the USA's rate appears to have remained fairly constant based on data supplied to the ILO" (though methodological caution applies). In this 2004 report, "Sweden and the UK [had] the lowest fatality rates" (p. 24).

The last two studies (one prepared by the UK HSE, the other by the US BLS) are both from 2014 and can be used to complement each other. The BLS's Wiatrowski and Janocha (2014) compare aggregated European (EU) OSH data with US one, while the HSE looks at the UK in comparison with other major EU economies. The US rate appears to be somewhat higher than the EU's in aggregate, and sometimes far higher at the sector level (Wiatrowski and Janocha 2014, p. 3). Within the EU, the HSE study (p. 2) shows that "the UK consistently has one of the lowest rates of fatal injury across the EU. In 2011 the standardised rate was 0.74 per 100 000 workers, which compares favourably with other large economies such as France (2.74 per 100 000 workers), Germany (0.94 per 100 000 workers), Italy (1.5 per 100 000 workers) and Spain (2.16 per 100 000 workers)".[673] Indeed, the UK's rate was roughly half the EU-15 average for most of the past decade and more.

Considering this data, it appears that both Australia (which has a fatality rate that in most recent years appears to be 40 to 50% lower than the US) and the UK (which has less than a third of the US's fatality rate) perform significantly better on at least this most easily comparable (and telling) measure of OSH. Differences in economic structure are substantial, but even when corrected for, the difference still exists, in the same direction. Australia and the UK both have two of the "enforcement styles" that most emphasize responsiveness, flexibility, promotion of compliance and overall focus on achieving safety outcomes rather than registering and sanctioning each and every violation. As we have seen, by statute and by design, OSHA's practice is in sharp contrast to this.

It would be difficult to conclude, however, on the relative level of effectiveness of OSHA's practices without considering the number of staff it can mobilize, and the number of inspections they conduct. As it states on its own website, and considering the size of the US economy and labour force, "federal OSHA is a small

[673] As discussed in earlier sections, these are the rates *excluding* fatal occupational accidents which took place in transit or transport.

agency"[674] – with "approximately 2,200 inspectors responsible for the health and safety of 130 million workers[675]". Part of the 50 US States have inspections and enforcement delegated by OSHA to State-level authorities, under an agreed plan – other have direct OSHA inspections and enforcement, hence the 2,200 inspectors include "state partners". By comparison, as we have seen earlier, the total number of (full-time equivalent) staff in the UK HSE and in Local Authorities working on OSH issues made up around 1,900. Britain, however, only has (latest Eurostat data) around 29.5 million workers overall (including self-employed), out of which slightly over 25 million *employees* (the main focus of OSH inspections in Britain). Thus, the ratio of inspectors to employees would be approximately 4.5 times higher in Britain, were there no other agencies involved in the US. In fact, there are agencies and structures at the local level that may be involved in OSH and are not summarized in OSHA's numbers, and there are some other federal administrations (most notably the US Mine Safety and Health Administration, MSHA) involved. These agencies are, in some cases, responsible for a distinct set of workers, but taking them into account may significantly change the ratio. For instance, the MSHA's 1,000 staff[676] are responsible for at most the slightly over 700,000 workers in the "Mining, Quarrying, and Oil and Gas Extraction" sector[677].

Likewise, while OSHA and its state-level partners conducted in 2015 slightly less than 80,000 inspections (federal inspections: 35,820 – StatePlan inspections: 43,471 – OSHA "Commonly Used Statistics" data), there were inspections conducted by other agencies (such as MSHA). These amount to approximately 43,000 visits per year on average over 2011-2014[678], including all kinds of visits (spot inspections on one single topic, reactive inspections, planned/regular inspections, information-focused visits etc.). Even accounting for such additional inspectors and inspections, however, it remains that their numbers appear low compared to prevailing levels in major EU countries, and even to Britain (where they have steadily decreased, and are far lower, as we have seen, than in Germany), at least considering the far larger working population. Indeed, a rough estimate would suggest that OSH inspections are at least 4 times less frequent in the US than in Britain, pro-rated to the working population. Things can look different, however, if we consider another unit of analysis, i.e. the number of active businesses (or of business establishments), and particularly those that are "above micro-size", since OSH inspections mostly focus on those where a significant number of workers is employed. Indeed, the US have, for a number of reasons, a very different enterprise structure from Europe, and a far smaller share of SMEs among the total number of businesses.[679] Britain has as of 2015 more than 5.2 million businesses, compared to 5.77 million in the US[680]. On this basis, the difference in the number of inspections per business would be negligible. If one considers only the businesses with at least 10 employees, the difference is however far larger (1.2 million in the US, around 235,000 in Britain). Thus, the ratio between the two countries will be very different depending on what is measured: inspections per workers, inspections per businesses, or inspections per businesses *above a certain size*. Overall, the total number of inspections in the US and Britain is relatively similar (around 120,000 if considering both OSHA and MSHA, and both HSE and

---

[674] Quote and data (including in next paragraph) from the "Commonly Used Statistics" of the OSHA website – available at: https://www.osha.gov/oshstats/commonstats.html

[675] Out of a total labour force of close to 160 million (see US Bureau of Labor Statistics data available at: http://www.bls.gov/news.release/empsit.t01.htm)

[676] Staffing numbers could not be found on the MSHA's website (http://www.msha.gov/) and were therefore obtained from the Wikipedia article: https://en.wikipedia.org/wiki/Mine_Safety_and_Health_Administration

[677] And in fact the MSHA is *not* responsible for all workers in this sector – data from the US Bureau of Labor Statistics available at: http://www.bls.gov/iag/tgs/iag21.htm

[678] MSHA does not publish anywhere a consolidated report on its number of visits. We had to extract this by analyzing the bulk data on all MSHA visits since the agency's creation, available at: http://ogesdw.dol.gov/views/data_catalogs.php (select "MSHA data" and then on next page "MSHA inspections"). Between 2011 and 2014, the total number of visits (all kinds) ranged from 41,174 to 46,366.

[679] See successive editions of the OECD's *SME Outlook* for details on this (the latest *SME Outlook* dates from 2015 but findings on this structural difference are still valid).

[680] US Census data for 2013 – see: http://www.census.gov/econ/susb/

LAs respectively, in most recent years) – but the ratio of working population (US/GB) is around 4.4, and the ratio of businesses with 5 employees or more is similar (4.44).

Thus, because on most measures OSH inspections in the US are roughly 4.4 times less frequent than even in Britain (itself a "low coverage" country by European standards), it is difficult to make any conclusions on the relative effectiveness of different inspection approaches between the US and countries with more risk-proportional, compliance-promoting countries. This has not prevented the fatal accidents rate to decrease in the US[681], but it has remained (as we have seen) significantly higher than in comparator countries. The fatal injuries' rate decrease was also slower than e.g. in Britain, where the HSE's figures put it at 84% between 1974 and 2015[682] (but of course the HSE had more staff, not only different methods).

What conclusion, if any, can we draw of this? Unfortunately, the existing studies and data tell us little about the relative effectiveness of US OSHA's approach compared to the British HSE – and tell us nothing about possible differences within the US (between federal OSHA and state partners, for instance). This also tells us very little about the question of focus and targeting. From Scholz and Gray (1990), we know that OSHA primarily targets larger firms. From OSHA's own website, we know that inspections heavily rely on "reactive" scheduling, while also incorporating an element of risk-based targeting.[683] The existence of a distinct administration (MSHA) also means that a very substantial share of the total OSH inspecting workforce is looking at a very small sub-set of the working population (admittedly, one that works in a high risk sector – but the institutional separation means that there can be no reallocation based on evolving risks). There are also risk factors that are US-specific, and which the action of OSHA inspectors is unable to affect. As US Bureau of Labor Statistics data shows[684], nearly 10% of fatal occupational injuries were homicides – whereas the number in Britain is nearly negligible.

There is, however, one area where it is possible that OSHA's approach, as mandated by its statute, may have had specific (negative) effects – it is in reinforcing political polarization around the agency's activities. While it is likely that (given the specific political climate in the US) *some* political polarization would have been present, the contrast with the US FDA (which elicits significantly less opposition) suggests that there may be some element of reaction to what Bardach and Kagan (1982) called "regulatory unreasonableness[685]". Since the low level of OSHA's staffing (and the resulting low level of inspections) are partly a factor of political opposition by the Republican party to any increase in OSHA's funding, and since hostility to OSHA among businesses was likely reinforced by the agency's "enforcement style", there may be a negative feedback loop between its approach and its effectiveness, not directly but mediated through its effect on the political and social acceptance of the agency's actions.

Overall, this short glance cast at the US situation and US-focused studies has raised more questions than it has yielded answers. Possibly, the decrease in fatal injuries would have happened regardless of regulatory interventions, given technological, economic, managerial and social change. The sharper decrease in the UK may reflect more resources, better methods, a different context – or all of the above. To our mind, this all strengthens the case to start by making more systematic comparisons – of resources, activities, methods and outcomes – before "drilling down" into statistics-based analytical work. This way, we may be better able to ask the right questions.

---

[681] By over 2/3 since 1970 and OSHA's creation – see: https://www.osha.gov/oshstats/commonstats.html

[682] See "historical picture" in the HSE statistics section: http://www.hse.gov.uk/Statistics/history/index.htm

[683] See "Inspections Fact Sheet" available at: https://www.osha.gov/OshDoc/data_General_Facts/factsheet-inspections.pdf

[684] See set of charts on fatal occupational injuries at: http://www.bls.gov/iif/oshwc/cfoi/cfch0012.pdf and a summary article on this topic (which has been the object of increasing public discussions) at: http://www.vox.com/2014/9/14/6139883/how-americans-die-on-the-job-in-5-charts

[685] And precisely OSHA examples formed a large part of Bardach and Kagan's book.

b. Real correlations or "noise" – assessing the quality of available data

We indicated above that there were two key reasons (apart from, obviously, limited time and resources) that we did not engage into systematic statistical analysis of data (looking for correlations in particular): inconclusive evidence from studies based on these methods (with conflicting findings) was the first, that we discussed in the previous section. The second was significant issues with data quality and reliability in many datasets – issues that become rapidly overwhelming once one attempts to correlated several insufficiently reliable data points, or to combine several insufficiently reliable sources.

Even in the OSHA study by Scholz and Gray (1990) that we have just discussed, not all data points are fully reliable. The records of inspections are most probably correct, but there is no absolute certainty that additional visits have not happened that were not recorded – while this may be insignificant in the case of OSHA, this is far from being negligible in other countries or institutions. Assessments of safety and health effects rely on reporting of injuries, diseases etc. that are inherently far less reliable, as we have discussed. As soon as a study seeks to take into account not only such relatively "objective" indicators but also more "qualitative" ones such as enforcement style, or some "hidden" ones such as corruption (broadly defined), or even "objective but hard to measure" such as burden – then the data becomes considerably less reliable. As a result, we would argue that studies that attempt to perform statistical analysis and establish correlations (let alone causalities) from such datasets tend to err, not because of their methodology, but because they apply what may be sound methods to profoundly unsound (or at least insufficiently reliable) data.

i. The difficulty of measuring corruption

Let us consider one simple example. At the onset of this research, we were hoping to include considerations of links between certain inspection and enforcement systems and practices on the one hand, and corruption prevalence in inspections on the other. Indeed, inspection and enforcement power can be abused by those who hold or oversee it, regardless of where orders come from, of what the law actually prescribes, of the existing safeguards etc. Such a possibility will always exist, and it would have been highly interesting to see whether certain systems seemed less corruption-prone than others. As we have discussed in earlier sections, there are reasons to believe for instance that performance management for inspection agencies based on public welfare outcomes decreases incentives for a particular form of "institutional corruption" whereby the agency tends to have an interest to find as many violations as possible to bolster its performance rating or its income (if they are linked to the number of inspection visits and of violations sanctioned). Systematically investigating this as well as other hypotheses on corruption would have required, however, some data that we could trust. Unfortunately, corruption is a very difficult phenomenon to *measure* (though it is easy to know it exists from anecdotal evidence). Indeed, the actual prevalence of corrupt behaviour is inherently difficult to measure and track. Corruption is by nature hidden, and most victims will be reluctant to report it for fear of reprisals - and in some cases the "victims" may in fact be rather willing to engage in corrupt behaviour, because it may be easier and less costly than compliance. Thus, it is very difficult to find adequate and reliable measures of corruption, and what data exists inevitably suffers from a number of limitations, regardless of the apparent precision given by scores and indices

Going further, establishing any causality between certain inspection regimes' features and corruption is even more problematic, at least if we look at demonstrating causality in a statistically- and quantitatively-grounded way. That would require not only assuming that responses to corruption-related questions were sincere (or at

least that under-reporting bias was constant in different jurisdictions and cases, which is highly unlikely), but also having databases which have been adequately checked for consistency and quality.

Considering the existing surveys which underpin the various indices on corruption (or exist as stand-alone), there are at least three different kinds of questions aiming at assessing corruption:

- Direct questions on whether the respondent has had to pay a bribe (or engage in corruption in any other way, e.g. gifts etc.), either in general to "facilitate" relations with authorities, or in direct relation with a given procedure (e.g. in our case an inspection);
- Indirect questions on whether the respondent considers that corruption is "prevalent" or "common" in a given situation (or in general in relations between enterprises and the state, or citizens and the state);
- Qualitative, rating questions, where respondents are asked to indicate how "severe" corruption is, or how much of a burden it is for their business (or businesses in general), or how "important" they consider corruption to be as a problem, etc.

Unfortunately, all three of these types of questions have shortcomings. The third category is the most obviously problematic: ratings are highly subjective, and depend on expectations, prevailing behaviours, existence or absence of comparison points, etc. They rarely correspond to what a data-driven analysis (e.g. growth or productivity factors) would indicate. They also rarely reflect the actual differences in prevalence of corruption, as examples in the following table will show. Possibly the only advantage of such questions is that answering them may be considered less "dangerous" by respondents, and so they may be relatively more open. But there are clearly major downsides in terms of reliability, particularly when trying to establish fine differences (i.e. between relatively similar countries, and not between "worst case" and "best case").

The first and second types of questions are somewhat similar, and only differ in that in the second case the question is asked more generally, and not necessarily in respect to the respondent. The second type of question is thus less precise (and, if we assume full truthfulness of replies, theoretically less reliable), and it is not possible (or at least more problematic - it depends on the exact wording of the question, as it may be restricted to "firms similar to yours") to use it to correlate corrupt behaviours with specific procedures, types of business, sectors etc.

By contrast, the first type of question is in principle the best one: it is unambiguous, precise, specific. It lends itself perfectly to quantitative analysis of any kind. The problem, however, is that available evidence suggests it is rarely fully honestly or truthfully answered (for reasons that are easy to comprehend), and also that its very precision lends itself to answers that are not "technically" false but in fact "hide" the reality. The first problem is linked to fear: in situations where corruption is prevalent, respondents (business, citizens) will frequently (but not always) fear for their safety if they report information that may be seen to be critical of the regime, or damaging for power holders. Since not all respondents will have full confidence in the strict confidentiality of the survey (whatever the assurances enumerators give them), a proportion of them will not respond truly - and this proportion varies, which makes any assumption to correct this error highly problematic, at best. The second problem is linked to the diversity of corruption - direct bribes paid during an administrative procedure or inspection are not the only form corruption takes, far from it. People may give gifts or payments at other times, support a higher official's child studying abroad, rely on friends or relatives' support to influence administrative decisions, etc. All these are *in fact* corruption of the proper working of the law, regulations, administration - but are not "making a payment to the official". Experience unfortunately suggests that most surveys do not take this issue adequately into account. And most reports analysing corruption-related questions tend to look for correlations and trends while taking all data at face value.

Formulation of questions, and truthfulness or accuracy of answers, are not the only issues affecting negatively data quality on corruption. The main problems that affect some, if not all, surveys are sample size and structure, and quality control and data cleaning. Since statistically representative surveys are a costly exercise, most such surveys are based on sample sizes that are set at the smallest possible size to produce statistically significant results on some key questions. That does not mean, however, because number of respondents may be lower for some questions, that the results are significant on *all* questions. And it is even less ensured that *variations* between different years or between countries can be measured accurately. In fact, such variations are often above the margin of error - and, contrary to a frequent situation with political polls, there are no additional measurements over several days or weeks to confirm or infirm the trends. Again, most analytical reports tend to avoid considering this issue too closely, and indeed it may not affect general conclusions (regressions over a number of countries, regional trends) too strongly, but it does preclude very precise comparisons between years or countries.

Quality control and data cleaning are possibly even more serious issues. Many surveys covering corruption issues are implemented across several countries, some of them in most of the world, which is essential to allow for comparisons. At the same time, budget constraints generally mean that the resources available for quality control (of translations of questions, of enumerators' training, of interview practices, of data entry and consolidation etc.) are limited. Practices also vary in terms of data cleaning (verification of outliers, possible exclusion of extreme outliers that cannot be explained by available knowledge or verification, etc.). This can result in aberrations being recorded as correct data, and this may be extremely difficult to spot. As a result, when data appears strongly at odds with "expert opinion", i.e. knowledge of the country based on first-hand experience and feedback from a number of direct personal interactions, it is not always clear whether this is due to the data being right and the "expert opinion" being based on biased experience, or whether the "expert opinion" is rather correct and the data happens to be of poor quality.

ii. Methodological and implementation limitations with available surveys

We have made use, in the section on the practical experience of post-Soviet and post-Communist countries, of business surveys conducted by the World Bank Group. Having been directly involved in the preparation and implementation of 5 of these, and indirectly in the supervision or analysis of many others, we are well aware of both the strengths and the limitations of these surveys. In terms of data quality overall, these surveys have benefited from large sample sizes, and strong supervision (not only by the survey firms, but by the World Bank Group teams in country), and extensive efforts to clean data by verifying consistency and cross-checking or eliminating outliers. When it comes to corruption data, however, there have repeatedly been serious problems that the survey teams were able to observe as respondent data came in. Unaccountably (at first glance), the percentage of respondents answering that corrupt practices had taken place during inspections (or other regulatory procedures) could suddenly drop from one survey to the next, whereas first-hand observations and numerous off-the-record interviews suggested this was by no means the case in reality. When attempting to cross-check the data in Tajikistan (where such a drop occurred between the 2003 and 2006 surveys), and following additional phone interviews with several hundred respondents, it became clear that the apparent drop was illusory. Many respondents did not respond "yes" because corruption took place *outside* of the inspection procedure itself (they used relations, or money, or gifts to ensure positive regulatory results – but did so *pre-emptively* and not during a particular procedure). Others were simply not convinced by assurances of confidentiality, and feared to answer truthfully. We have similarly observed very low levels of answers on corruption-related questions in regimes that were strongly authoritarian. A further limitation of these surveys (and the reason we have not tried to use them to study other possible correlations) is that they cover only a limited set of countries (several post-Soviet republics, and Mongolia), with a limited number of years (frequent surveys for some countries, like Ukraine, but only a few for others, like Mongolia or Georgia).

By contrast, another group of surveys, called worldwide the "World Bank Enterprise Surveys[686]", and known as Business Environment and Enterprise Performance Surveys (BEEPS) in Eastern Europe and Central Asia (where they are run jointly by the European Bank for Reconstruction and Development and the World Bank Group – the EBRD taking the lead[687]), offers global coverage (including a number of EU/OECD countries), and regular iterations. These surveys have been used for a very large number of research pieces[688], including many focusing specifically on regulatory issues, and on corruption.[689] Some research has in fact tried and use the large coverage of the Enterprise Surveys dataset to assess whether patterns of corruption under-reporting can be found, and indeed found under-reporting to be higher under more authoritarian regimes (Jensen, Li and Rahman 2007) – but most others have looked at corruption data as being reliable, and used it to investigate correlations with other variables.

Some of these papers reach results that are "challenging" or even difficult to understand. Knack and Kisunko (2011) thus report at face-value changes in use of bribes, and end up with (relatively free) Kyrgyzstan having far higher levels than (far more authoritarian) Azerbaijan or Tajikistan – something that is clearly difficult to reconcile with first hand observations in these countries. The answers to different questions were also seemingly difficult to reconcile for some countries (e.g. Kosovo, where many respondents both answered that bribes were rare, and that corruption was a major problem), leading the authors to elaborate complex interpretative theories, without first considering whether data quality issues may play a role. For their part, Blagojevic and Damijan conclude among other points that "results suggest that foreign owned firms are more likely to engage in informal payments" (p. 20), not considering whether they may not in fact just be more forthcoming with their answers (something that would be quite likely in our experience). They also write that most negative effects of corruption on firm performance "dissipate after 2004, indicating the general improvements in the business environment" (*ibid.*). This *might* be true but could also indicate a decrease in the level of corruption reporting by respondents (something we have repeatedly observed).

To conclude on this issue, and demonstrate more clearly why we decided *not* to attempt any systematic statistical analysis (least of all on corruption, but not on other issues either), let us look more closely at some examples from Enterprise Surveys data.

Enterprise Surveys have been conducted every few years since 2002, relying on face-to-face interviews of business operators in up to 135 different countries (the exact number can vary between different iterations). The methodology has been fully harmonised since 2005-06. The survey instrument covers a number of aspects of business operation - legal and regulatory environment, infrastructure, technology and skills, access to finance etc. The sections that are of interest to us here are the ones on "Corruption" and "Regulations and Taxes".

The WBG Enterprise Surveys (and EBRD-WBG BEEPS) have the advantage of covering a large sample of countries, including some developed (OECD, EU) ones, which allows to compare very different regulatory regimes. They have a range of questions covering corruption, but have no questions directly on inspections. They have questions on the regulatory environment e.g. licensing and construction permits (but these are rather bad proxies for inspections, as one part of the regulatory system is not always a predictor of another), and a question on "interactions with tax authorities" that *includes* tax inspections, but is not limited to them (visits to the tax authorities, e.g. to file or pay taxes, are also included). Thus, this survey can mostly be useful

[686] See website: http://www.enterprisesurveys.org/
[687] See website: http://ebrd-beeps.com/
[688] See databases of research done using these surveys – on World Bank website: http://www.enterprisesurveys.org/research - on EBRD website: http://ebrd-beeps.com/research/
[689] To name but a few examples: Knack and Kisunko 2011, Blagojevic and Damijan 2012, Denisova-Schmidt and Huber 2014

to provide data points on corruption for a large number of countries, though if no other data is available on a given country the "interactions with tax authorities" question can be used as an imperfect proxy.

There are however very significant methodological issues with the Enterprise Surveys. The first is the extensive use of "qualitative", "opinion" questions, asking respondents to rate the importance or severity of a constraint. We provide below examples of how such questions can provide results that simply do *not* reflect the real condition in a given country. The second is the relative weakness of the control on survey implementation, and also the lack of attention to whether questions are actually phrased in a way that *can* give meaningful results (i.e. if the questions simply *make sense* in a given country, or not). The third is that sample size is rather small, and that the sample does not include micro-businesses and individual entrepreneurs (which, typically, are quite affected by "petty corruption") - thus, the statistical reliability of the data (in terms of providing a reliable picture of the whole economy of a country) is not fully assured. The table below illustrates the problems with corruption-related data in the Enterprise Surveys.

Corruption data and its inconsistencies – WBG Enterprise Surveys / BEEPS

Economy	Year	Bribery incidence (percent of firms experiencing at least one bribe payment request)	Percent of firms expected to give gifts in meetings with tax officials	Percent of firms expected to give gifts to get an operating license	Percent of firms expected to give gifts to get a construction permit	Percent of firms expected to give gifts to public officials "to get things done"	Percent of firms identifying corruption as a major constraint
Armenia	2002	...	34,0	...	...	25,0	15,3
Armenia	2005	...	70,5	...	...	25,9	21,4
Armenia	2009	15,5	13,3	11,3	21,7	16,0	39,6
Germany	2005	...	14,8	...	...	...	3,9
Kyrgyz Republic	2002	...	79,1	...	...	63,3	17,5
Kyrgyz Republic	2003	...	49,5	25,0	56,3	77,4	50,5
Kyrgyz Republic	2005	...	83,8	...	...	68,3	34,3
Kyrgyz Republic	2009	42,5	39,0	25,7	56,3	47,8	58,9

Kosovo	2009	9,0	7,7	5,8	5,7	7,5	73,4
Lithuania	2002	...	29,1	...	...	49,1	18,7
Lithuania	2004	...	3,8	8,2	5,6	...	27,6
Lithuania	2005	...	35,4	...	...	43,4	16,6
Lithuania	2009	7,2	3,4	5,4	32,4	10,7	38,6
Latvia	2002	...	37,1	...	...	48,4	15,3
Latvia	2005	...	30,0	...	...	29,8	12,8
Latvia	2009	8,9	4,4	0,0	13,2	13,4	33,9
Slovenia	2002	...	15,8	...	...	15,0	4,9
Slovenia	2005	...	13,9	...	...	12,2	4,8
Slovenia	2009	2,3	0,0	0,0	3,6	5,8	9,8
Tajikistan	2002	...	84,4	...	...	68,5	18,9
Tajikistan	2003	...	44,4	84,8	86,7	55,3	23,5
Tajikistan	2005	...	71,5	...	...	51,1	20,9
Tajikistan	2008	37,9	33,0	38,5	42,3	44,6	37,8

Source – World Bank Group Enterprise Surveys for the years indicated

Through this table and subsequent ones, we have tried to show some of the problems with corruption-related data in Enterprise Surveys. To this aim, we have taken a few countries from the former Soviet Union (on which other data sources and information are available so as to cross-check the survey data), a few recent EU Member States (as "transition" comparators), and some Western European countries for reference[690].

A number of data points appear to be particularly problematic, which we outline below.

[690] Limiting ourselves to those available in the database, which does not cover all of Western Europe, and covers these countries in some years only, as they are only there as reference points and not as the survey's main focus.

First, in Kosovo, nearly ¾ of respondents rate corruption "a major constraint" but less than 10% respond positively on any of the other corruption-related questions (which are "objective" questions, i.e. of the type "did something – like paying a bribe – take place or not"). This suggests that answers to the "major constraint" may not be fully meaningful.

Second, in Germany, nearly 15% of respondents indicated that "giving gifts when meeting tax officials" was expected. This appears not to be supported by any other data, and is clearly not particularly consistent with what is usually known of the experience of firms in that country [note that the same indicator in Slovenia for that year was 0%]. It is unclear how reliable the answers are.

Third, in many cases, variations from one year to the other appear difficult to reconcile with any plausible explanation. To the author's best knowledge (with direct work experience in this field in these countries), there exists no particular set of reforms, changes or events which could explain these. For instance, the percentage of firms in Armenia reporting that they were expected to give gifts to tax officials first doubling (2002-2005) then decreasing 80%. This is compounded by internal inconsistency, as between 2005 and 2009 the percentage of respondents rating corruption "a major constraint" nearly doubled (suggesting an opposite trend). The same indicator on gifts to tax officials in Kyrgyzstan decreasing nearly 40% from 2002 to 2003, then nearly doubling, then halving again. In addition, still in Kyrgyzstan, we see the indicator "corruption as a major constraint" fluctuating wildly (nearly tripling first, then 30% down, then again up 75%) [note: major political changes did take place in Kyrgyzstan in this period, but not in 2002-2003, and in any case changes in tax administration and public administration overall were far from radical]. Likewise, in Lithuania we observe variations by a factor of 10 on the "gifts to tax officials" indicator between two consecutive surveys, repeatedly – and "wild" (if not quite as large) swings on this same indicator in Tajikistan.

In addition, internal inconsistencies are numerous – indicators on "gift to tax officials", "gifts to public officials" and "corruption as a major constraint" frequently move in opposite directions from one survey to the next (particularly in Tajikistan). Furthermore, some of the trends, even not considering all of the above, appear not to reflect reality as experienced and reflected through in depth interviews, daily experience etc. in various countries – and likewise for the relative level of indicators across different countries. Indeed, data from the most recent survey suggests that corruption in administrative procedures is higher in Kyrgyzstan than in Tajikistan, whereas most other evidence (country-specific surveys, in depth interviews, etc.) suggest that corruption in Tajikistan is *at least* as high as in Kyrgyzstan (and probably higher). While this does not in and of itself prove that the data is incorrect, when combined with the above, it certainly suggests that it could be the case. In a similar fashion, absolute levels in percentage of firms expected to give gifts to get a license or permit, even though high in Tajikistan or Armenia, seems far lower than what most in depth interviews and direct experience would suggest. This, again, raises additional doubts – particularly when combined with all the issues outlined above.

Taking a closer look at other modules of the Enterprise Surveys raises similar problems. These problems are worth considering because the Enterprise Surveys are, in fact, not worse than many other instruments routinely used for analytical reports, studies and papers – and possibly better than many. These flaws suggest that far too much confidence is routinely given to such quantitative data, resulting in researchers and analysis building complex models and testing causality, all on foundations that appear quite flimsy upon closer inspection.

Regulatory issues data and its problems – WBG Enterprise Surveys

Economy	Year	Senior management time spent dealing with the requirements of government regulation (%)	Number of visits or required meetings with tax officials	Days to obtain an operating license	Days to obtain a construction -related permit	Percent of firms identifying tax administration as a major constraint	Percent of firms identifying business licensing and permits as a major constraint
Armenia	2002	1,9	...	...	...	37,1	7,6
Armenia	2005	3,1	2,8	...	...	46,7	16,8
Armenia	2009	10,3	2,1	20,0	26,3	21,1	5,6
Germany	2005	1,2	1,3	...	...	23,2	4,0
Spain	2005	0,8	1,5	...	...	12,8	12,9
Estonia	2002	2,1	...	...	...	5,1	11,8
Estonia	2005	2,1	0,3	...	...	3,6	3,6
Estonia	2009	5,5	0,4	8,3	29,0	3,1	3,4
Kyrgyz Republic	2002	5,1	...	...	...	23,7	9,4
Kyrgyz Republic	2003	6,6	16,6	43,9	115,6	52,5	13,4
Kyrgyz Republic	2005	5,1	3,5	...	...	34,2	11,8
Kyrgyz Republic	2009	4,9	2,1	18,0	64,6	31,6	16,3
Kosovo	2009	9,8	4,5	18,8	47,7	10,4	7,5
Lithuania	2002	5,4	...	...	...	22,7	9,4
Lithuania	2004	25,9	9,5	55,5	63,4	36,8	13,4

Lithuania	2005	4,5	2,2	...	...	15,0	9,3
Lithuania	2009	9,3	0,8	65,5	44,6	28,0	23,4
Latvia	2002	5,8	...	...	...	28,4	9,8
Latvia	2005	3,2	1,4	...	...	27,9	8,4
Latvia	2009	9,7	1,5	11,5	70,0	25,0	14,4
Slovenia	2002	3,8	...	...	...	7,1	2,4
Slovenia	2005	2,4	0,4	...	...	15,4	1,6
Slovenia	2009	7,3	0,3	56,0	117,0	5,7	4,6
Tajikistan	2002	5,8	...	...	...	25,2	15,0
Tajikistan	2003	1,5	6,2	15,3	17,2	9,5	9,0
Tajikistan	2005	4,2	2,5	...	...	22,9	14,3
Tajikistan	2008	11,7	1,4	22,6	62,0	17,0	16,6

Source – World Bank Group Enterprise Surveys for the years indicated

As we can see looking at this table, data on administrative and regulatory procedures is quite problematic as well. We summarize below some of the main difficulties.

First, one of the indicators is “days to obtain an operating license” – however there is no clear definition of what an “operating license” is. The Enterprise Survey has some of its origins in work on Latin America, and “operating licenses” are quite frequent in that region. There, they are an additional approval that is mostly given out by municipalities, and comes on top of the business registration (and of tax registration too). They are “universal” in the sense that they apply to all businesses, regardless of sector etc. The Enterprise Survey team decided that from 2005-06 they would have a fully harmonised (unified) methodology worldwide, and not only kept this indicator for all regions, but adopted guidelines that prohibit any clarification of indicators by enumerators, to avoid any variation between countries (at least this is the intent, if not the result). The problem (and this is not a minor one) is that “operating licenses” of this kind do not exist in all parts of the world (though they exist in much of Africa, in addition to Latin America). Notably, they rarely exist in Europe (both inside and outside of the EU), with the exception of Greece (where they are not fully universal, but close to it). Nor do they exist in the Former Soviet Union, for instance. This does not prevent the surveys from asking the questions, recording responses, and then does not prevent research articles or reports from analysing this data. It is likely that either translators in each country will give a different meaning (one that “seems to make sense”), or each respondent will answer based on whatever s/he assumes to be meant (probably taking whichever license he has recently obtained for a particular activity – but not all businesses obtain licenses, so

many respondents must answer based on their experience with other procedures). This introduces major scope for data errors.

Second, several countries show massive increases in the "senior management time spent dealing with government regulation" over time, particularly in the 2009 survey. Again, this happens without any known event or change which could be a credible explanation. On the contrary, in most countries affected by this increase, important regulatory simplification reforms took place. Even assuming very imperfect implementation of these reforms, a radical worsening of this kind is surprising, and is more likely to indicate low reliability of the indicator. Indeed, the question is difficult to answer reliably for many respondents. It could also reflect the effects of an unreliably low sample size. For instance Armenia sees this indicator triple from 2005 to 2009, and Tajikistan is also close to tripling, as well as Latvia, Slovenia. In Lithuania, there is first an increase (by a factor of 5) from 2002 to 2004, then a decrease (by the same factor), then a doubling. Such variations suggest that the indicator is essentially random (i.e. respondents answer "whatever comes to mind").

Finally, "number of visits or required meetings with tax officials" is an indicator that could be useful for our purpose (as a partial proxy for "inspections"). Unfortunately, the quality of this indicator is doubtful. For Tajikistan, it shows a very rapid decrease that is not confirmed through any of the other available surveys (which have larger sample sizes and a narrower focus on the topic). In addition, the reported number of interactions in Tajikistan is several times lower than only the number of inspections (as measured by the more detailed World Bank Group surveys we presented earlier, which had a far larger sample size and more "intensive" quality control), without even counting the many visits to the tax authorities needed to file for many taxes there. In Kyrgzystan, the scope of the decrease appears far larger than what other available data, as well as the contents of reforms, would suggest.

A final look at another part of the Enterprise Survey data (infrastructure, including access and quality) will allow us to further justify our refusal to engage in complex statistical analysis on the basis of any of this data, and to consider all "subjective" questions (asking respondents to rate the importance of a problem) with the utmost caution. Looking at the questions on infrastructure indeed allows to compare such ratings with easily observable, objective reality on infrastructure quality in different countries.

Enterprise survey data on infrastructure – reality and opinion

Economy	Year	Days to obtain an electrical connection (upon application)	Percent of firms identifying electricity as a major constraint	Percent of firms identifying transportation as a major constraint
Armenia	2002	2,5	13,1	10,4
Armenia	2005	2,5	3,1	10,2
Armenia	2009	16,3	24,9	26,3

Germany	2005	3,2	1,0	1,6
Spain	2005	9,3	8,3	10,8
Kyrgyz Republic	2002	3,5	1,7	1,7
Kyrgyz Republic	2003	22,1	4,0	3,0
Kyrgyz Republic	2005	14,1	1,3	1,3
Kyrgyz Republic	2009	25,2	58,0	30,9
Slovenia	2002	9,6	0,8	0,0
Slovenia	2005	12,7	2,3	0,8
Slovenia	2009	60,8	23,3	12,3
Tajikistan	2002	12,5	15,0	3,8
Tajikistan	2003	11,6	22,1	5,2
Tajikistan	2005	5,3	10,8	2,2
Tajikistan	2008	28,3	39,2	22,9

Source – World Bank Group Enterprise Surveys for the years indicated

Once again, the data appears plagued by two combined problems: unreliability of data in terms of internal consistency (suggesting problems with data quality), and unreliability of "identification as a major constraint" questions as an indicator of the actual situation for businesses. Looking at a few examples makes it clear.

As a first problem, we see wild swings in the days needed to obtain an electrical connection in Tajikistan from one year to the next, again not related to any known change in the system and practices, suggest that the number of respondents is too low, or some outliers or data entry errors cause the average (mean) value to increase or decrease without reflecting the "actual average" in reality. The massive increases in days needed for such a connection in Armenia and Slovenia also appear difficult to explain. Since systematic checks of outliers are not necessarily undertaken by contractors, these may be enumerators' errors, data entry errors, or whatever else.

A second illustration can be seen on transportation. In practical terms, transportation in Tajikistan is a major issue (lack of paved roads and/or horrendous conditions of said roads, very limited rail links to the outside, rare international flights etc.) – but only 2-5% of respondents in 2002-2005 rated it as a "major constraint". While real improvements (however limited) took place after 2005, the percentage suddenly jumped by a factor of 10 (to nearly 23%). This suggests that people respond to such questions based on the salience of various issues for them at a given time, their expectations, their points of comparison (or lack thereof), possibly the

order of questions, etc. A roughly similar comment can be made for data on Kyrgyzstan and Armenia, which clearly understates the extent to which transportation is a problem in these countries. In the same vein, in 2005, nearly 11% of respondents in Spain rated transportation a major constraint – which is far more than in landlocked Tajikistan and Kyrgyzstan, more even than in Armenia. While this may correspond in some cases to different propensities to trade beyond an enterprise's home city (if no trade is undertaken beyond a few kilometres, transportation may indeed not be a "constraint"), it is nonetheless clear that these responses do not in any meaningful way reflect the objective situation in terms of quality of transport infrastructure, transportation times etc. As a result, such data is quite simply irrelevant when one tries to determine the "picture" for a given country in terms of constraints for economic development.

Considering these major flaws in available data (and the Enterprise Surveys are most likely not worse than most other "global" datasets – the easy accessibility of the detailed data just makes it easy to point out the problems), we have decided not to attempt any statistical analysis – regressions and any other tools pointing to correlation (or lack thereof), and more complex tools to try and isolate specific factors and causal links. We simply believe that such analysis would be close to meaningless considering the unreliability of the data, whether caused by poor data quality (insufficient samples, lack of data verification and cleaning etc.), excessively complex questions (to which respondents *cannot* reliably answer because they assume a level of knowledge they do not have, e.g. when asked about percentage of their time spent on something in a given year), or "inherently unreliable" questions (where people are asked to rate "major constraint", which will be answered based on salience of experience, which is not what is relevant to analyse the situation in the country).

We have, rather, chosen to limit ourselves to observing aggregate, country-level trends. We believe that the results thus achieved are clearly limited in terms of establishing correlations or causations, but that they are more solid in the sense that they are not built on fragile data. We have indeed chosen to rely only on surveys where we had a reasonably high level of confidence in the data quality, and within these surveys only on questions where the objectivity and straightforwardness of questions minimized the risk of bias or error. We have also taken official (inspectorates, Eurostat etc.) data, again selecting only those sources and indicators that were most reliable. In the way we used this data, we have avoided attempting to conduct statistical correlation or other analysis, but rather looked at high level, aggregate comparisons. Because they all leaned in the same direction, and in spite of the limitations of the approach, we hope these case studies have shed some light on the relative effects of different inspections and enforcement approaches, and indicated some directions for future research.