Economics

A Compliance Return method to evaluate different approaches to implementing regulations: the example of food hygiene standards

--Manuscript Draft--

Manuscript Number:	ECONJOURNAL-D-23-00157
Full Title:	A Compliance Return method to evaluate different approaches to implementing regulations: the example of food hygiene standards
Article Type:	Research Article
Keywords:	Evaluation; Approaches to implementing regulations; Food safety regulation; 'Compliance return'; Local government.
Manuscript Region of Origin:	UNITED KINGDOM
Abstract:	We investigate how different approaches to the implementation of regulations affect compliance by evaluating the implementation of food hygiene standards in four English authorities. We draw upon regulatory theory in general together with specific literature on food hygiene standards to advance a new ‘compliance return’ concept, which enables compliance outcomes and resource inputs to be considered jointly. We operationalise ‘compliance return’ as part of a low-cost methodology for evaluating the relative effectiveness of different approaches to implementing food hygiene standards. We find that different approaches to implementation have significant impacts on food policy outcomes. Specifically, implementation ranged from a greater emphasis on deterrence to a greater emphasis on cooperation, with cooperation resulting in greater compliance and deterrence requiring less resource input. When compliance level is considered alongside resource input, to assess the overall ‘compliance return’, a stronger case emerges for a deterrence approach. The food policy implications are two-fold: (i) implementation affects policy outcomes; and (ii) the measure of ‘compliance return’ developed and implemented in this article offers a low-cost approach to support the ex-post evaluation of implementation choices. The Compliance Return approach may be applicable more widely, especially where there is discretion regarding how regulations are to be implemented.
Manuscript Classifications:	8.5.4: Project Evaluation • Social Discount Rate; 11.3: Regulation and Business Law; 17.4.9: Agricultural Policy • Food Policy

A *Compliance Return* method to evaluate different approaches to implementing regulations: the example of food hygiene standards

Running title: Compliance return approach to implementing regulations

Abstract.

We investigate how different approaches to the implementation of regulations affect compliance by evaluating the implementation of food hygiene standards in four English authorities. We draw upon regulatory theory in general together with specific literature on food hygiene standards to advance a new 'compliance return' concept, which enables compliance outcomes and resource inputs to be considered jointly. We operationalise 'compliance return' as part of a low-cost methodology for evaluating the relative effectiveness of different approaches to implementing food hygiene standards. We find that different approaches to implementation have significant impacts on food policy outcomes. Specifically, implementation ranged from a greater emphasis on deterrence to a greater emphasis on cooperation, with cooperation resulting in greater compliance and deterrence requiring less resource input. When compliance level is considered alongside resource input, to assess the overall 'compliance return', a stronger case emerges for a deterrence approach. The food policy implications are two-fold: (i) implementation affects policy outcomes; and (ii) the measure of 'compliance return' developed and implemented in this article offers a low-cost approach to support the *ex-post* evaluation of implementation choices. The *Compliance Return* approach may be applicable more widely, especially where there is discretion regarding how regulations are to be implemented.

Key words: Evaluation; Approaches to implementing regulations; Food safety regulation; 'Compliance return'; Local government.

Highlights:

- How you implement food hygiene standards affects policy outcomes.
- Support-to-compliance (emphasizing cooperation) achieves marginally higher compliance.
- Compliance-to-support (emphasizing deterrence) achieves marginally lower compliance at lower cost.
- A new *compliance return* measure provides a low-cost method of evaluating implementation choices.

1. Introduction

This article addresses practitioner concerns regarding the implementation of UK food hygiene regulations for establishments providing food directly to consumers and contributes to the related scholarly literature on approaches to regulatory compliance. Practitioners – e.g., senior public health managers in local authorities – must not only implement food hygiene standards but must do so within cost constraints. This challenge has consequences when failure to comply affects public health (e.g., there are an estimated 2.4 million cases of foodborne illnesses each year in the UK, within which there are 380,000 cases of norovirus, with eating out and takeaways accounting for 37 percent and 26 percent of those cases; Food Standards Agency, 2020). Yet, especially since 2009-10, when financial crisis triggered "austerity" policies that severely constrained local government budgets, local officials must ensure compliance with limited resources (Regulating Futures Review, 2017).

The practical need to reconcile high standards of public health with low enforcement costs of doing so gave rise to the research reported in this article, which was commissioned by a consortium of four neighboring District Councils (local authorities; henceforth, LAs) – each responsible for implementing the national food hygiene regulations in its own district – in the English West Midlands. (To preserve anonymity, these four LAs are denoted LA1, LA2, LA3 and LA4.) To this end, we developed a 'compliance return' construct to (i) evaluate the efficacy of different approaches to the implementation of food safety regulations and (ii) demonstrate the terms of the tradeoff between food safety standards and the cost of securing compliance. Further, we develop and explain how our quantitative methodology can deliver these two objectives using the data available to managers, hence at low cost.

In pursuit of these objectives, we also contribute to understanding how implementation affects compliance. To assess this, we focus on two key compliance approaches, which, in their different guises, have deterrence and cooperation as opposing or complementary strategies (Ayres and Braithwaite, 1992; Bates et al., 2015; Braga et al., 2018; Burby and Paterson, 1993; Downs et al., 1996; Kuperan and Sutinan, 1998; Mackay et al., 2020; OECD, 2013, Rechtschaffen, 1998; Scholz,1984). Our contribution to this literature is to investigate, in the context of food hygiene, how even small changes along this continuum of regulatory approaches can have substantial impacts on both compliance outcomes and compliance costs.

To enable implementation practices and their outcomes to be compared across the four local authorities, we were granted access to a large longitudinal dataset recording the complete results from food hygiene inspections carried out by all four authorities over a 10-year period (fiscal years 2008/09 – 2017/18): i.e., 26,285 individual assessments of food hygiene and safety performance carried out during regulatory interventions at 11,294 establishments. These authorities also provided personnel data to enable us to consider the resources involved in the delivery of the enforcement service, which, when considered alongside levels of compliance, enabled a fuller assessment of the respective 'compliance returns' of the different approaches to implementation.

This study commences by positioning the investigation within the deterrence and cooperation literature and, following this, the literature on compliance within the food hygiene and safety area. We then introduce the construct of 'compliance return' and demonstrate its use in comparing different approaches to implementation along the deterrence-cooperation

continuum. This involves explaining the two essential constituents in the measure of 'compliance return': the compliance score; and the resource input by the regulator. Having considered the literature on compliance and the rationale for our 'compliance return' construct, we present the methods used to analyze the data, present and interpret the results of the analysis, and discuss the policy implications of these findings.

Regarding the wider applicability of our *Compliance Return* approach to evaluation, we conclude that it may have external validity in regulatory fields where (i) full cost-benefit analysis is not feasible, (ii) there is some discretion – and thus some variation – regarding implementation, and (iii) consistent records of inspection results are maintained across implementing authorities. In this article, we do not attempt to identify such regulatory fields. Although we cannot make statistically verifiable claims to external validity, we attempt to provide sufficient detail about the context and implementation of our approach for readers to make informed judgements as to the applicability of our evaluation methodology to their own circumstances.

2. Policy: modes of implementation

In this Section, we draw upon the literature to identify deterrence and cooperation as modes of implementing food hygiene and safety policy.

2.1. Theory (1): Using Deterrence and Cooperation to Secure Compliance

When the success of a regulation is dependent on the level of compliance achieved, and when compliance is not automatic, then the question of how compliance can be encouraged becomes a critical matter. The question concerns how compliance can be incentivized most effectively, with the focus typically being a penalty to deter non-compliance and/or support to help achieve compliance. These represent two styles of implementation and enforcement. The names given to these styles include punishment and persuasion (Braithwaitet, 1985), deterrence and cooperation (Rechtscaffen, 1998), and coercive and catalytic (Weske et al., 2018). Each enforcement style has different assumptions and actions. Deterrence assumes that individuals (as well as private and public organizations) make rational choices based on self-interest. This can lead to an unwillingness to comply, which, being a purposeful choice, is deemed to justify and require compulsion to force compliance (Becker, 1968; Cohen, 2000; Ehrlich, 1975; Kagan et al., 2003; Mintz, 1995; Markel, 2000, 2005; Nagin, 2013; Pearce and Tombs, 1997; Stigler, 1970). Compulsion can include warnings, fines, closure, and imprisonment. Cooperation assumes that those being regulated are willing to comply but lack the capacity (e.g. knowledge) to comply. Actions would then involve support to address the shortfalls in capacity (e.g. training). Cooperation as a normative approach was born from studies showing that persuasion, assistance, and collaboration was the preferred approach of regulators (Braithwaite, 1985; Cranston, 1979; Hawkins, 1984, 2002; Hutter, 1997), and from criticisms of deterrence theory (Andreen, 2007; Rechtschaffen and Markell, 2003; Scholz, 1997; Stoughton et al., 2001). These criticisms, which Scholz (1997) summarized in four key points, are that (i) corporations fail to act as fully informed utility maximizers, (ii) regulation fails to unambiguously define the required behavior, (iii) punishment is not a primary driver

of behavior, and (iv) officials do not have the capacity to detect and punish non-compliance. Together, these constitute the justification for a more cooperative focused approach.

Deterrence and cooperation should not however be seen as mutually exclusive. It is recognized that cooperation may not always work, and that coercion needs to be available, to be called upon if necessary (Ayres and Braithwaite, 1992; Braithwaite 1985, Burby and Paterson, 1993; Hawkins, 2002; Hutter, 1988; OECD, 2005; Rechtschaffen, 2004; Scholz, 1997; Stoughton et al., 2001). Integration rather than exclusion has therefore permeated the debate on deterrence and cooperation, with much of the detail being on when and where each should be used and their relative effectiveness (Andreoni et al., 2003; Chen et al., 2012; Glicksman and Earnhart, 2007; Scholz, 1984; Scholz and Gray, 1997). However, what remains unclear is the issue of effectiveness more broadly. In particular, whilst the most obvious criterion is the level of compliance achieved, this cannot be compliance at any cost, due to the expectation that the benefits from compliance should outweigh the costs incurred (e.g., the remit of Regulatory Policy Committee, created by the UK Government in 2009, focuses on the impact of regulatory decisions). Accordingly, to compare the *efficacy* of different approaches to regulatory compliance with food hygiene standards, we develop and apply our 'compliance return' construct, which jointly measures compliance and the costs to the regulatory authority of achieving compliance. Joint measurement enables policy makers to use existing data – hence, at minimal cost – to evaluate different approaches to implementation by exploring potential trade-offs between levels of compliance and the costs of achieving such levels.

Empirically, studies on effectiveness have tended to focus on *either* deterrence *or* cooperation, levels of compliance when using *either* deterrence *or* cooperation, or the effect of *either* deterrence *or* cooperation on the regulatee and/or the regulator. Examples include the explanation of provisional driver non-compliance using deterrence theory (Bates et al., 2015); cooperation being an incomplete explanation of compliance with international agreements (Downs et al., 1996); deterrence being more effective when accompanied with a moral obligation and social pressure for the regulation of Malaysian fishermen (Kuperan and Sutinan, 1998); cooperation through capacity building improving compliance with State environmental regulations, and deterrence working as well for simpler regulatory requirements (Burby and Paterson, 1993); a systematic review showing that deterrence produces a small reduction in gang violence, drug markets and repeat offending (Braga, 2018); that increased compliance occurs when firms participate in the compliance process (Malesky and Taussig, 2017); and that regulators will make trade-offs between deterrence and cooperation when seeking compliance (Gunningham, 2017). One study that compared deterrence and cooperation for the same regulated activity is Weske et al. (2018). They compared the use of deterrence and cooperation for regulation on quality and patient safety in hospitals, showing that cooperation increased the capacity of ward leaders to comply, subject to the caveat that this depended on the motivation of ward leaders, which was shown to vary, which in turn necessitated the use of deterrence when motivation was absent. Another study, by Earnhart and Glicksman (2015), investigated deterrence and cooperation for water discharge limits, concluding that cooperation resulted in better environmental management by the firm.

The paucity of studies assessing how regulation is carried out is a significant gap. Yet, the implementation literature is clear in pointing out the importance of identifying and assessing the implementation influencing variables in the securing of compliance (Khanna, 2021; Peters et al., 2013; Sabatier, 1986). These variables include the consistency and effectiveness of inspectors (Barnes et al., 2022), availability of implementation resources (Borraz et al., 2022; Luukkannen et al., 2018), and forms of cooperation (Buckley, 2015; Nevas et al., 2013). In our analysis, the implementation influencing variables, as previously alluded to, are the use of deterrence and the use of cooperation, and their implementation effectiveness as measured by compliance output per unit of resource input. Specifically, when using deterrence and cooperation for the same regulated task, there is an absence of empirical analysis of compliance levels alongside analysis of the efficacy in delivering those compliance levels. To address this deficiency, we analyse efficacy by (i) comparing compliance levels when using different approaches to regulation to secure compliance with food hygiene standards by establishments providing food directly to consumers, in conjunction with (ii) comparing the resources used in the delivery of these two approaches to securing compliance. Before this is investigated in detail, we review the literature on compliance with food hygiene and safety standards.

2.2. Theory (2): Food Hygiene and Safety Standards

Food hygiene and safety in England, Wales and Northern Ireland is overseen by the Food Standards Agency and in Scotland by Food Standards Scotland. The legislation the Food Standards Agency implements includes the Food Safety Act 1990, the Food Safety and Hygiene (England) Regulations 2013, Regulation (EC) 178/2002, Regulation 852/2004 and

other areas of law as detailed in the Food Law Code of Practice (England) 2017. Local authorities are responsible for implementation, which for England, Wales and Northern Ireland involves 329 local authorities. A large proportion of the establishments that are subject to the food hygiene and safety standards supply food directly to the consumer. The food hygiene and safety standards are operated through a food hygiene rating scheme, which involves separate ratings on 'food hygiene and safety procedures', 'structure', and 'confidence in management'. (Structural requirements concern the cleanliness and ease of cleaning of all surfaces together with the adequacy of washing facilities, ventilation, lighting and drainage; 'confidence in management' is a measure of confidence in the business's food safety management systems and likelihood of future compliance.) The rating is from 0 to 5: where 0 = urgent improvement necessary; 1 = major improvement necessary; 2 = improvement needed; 3 = generally satisfactory; 4 = good; and 5 = very good. The rating is compulsorily displayed at the entrance to establishments in Wales and Northern Ireland, and voluntarily displayed in England. Ratings are also publicly accessible via the Food Standards Agency website.

Studies on the use of food hygiene and safety standards, and the visual display of establishment performance, generally show a positive effect on food hygiene and safety. A study that looked at the use of inspection based 'grade cards' for hygiene performance for restaurants in Los Angeles County, showed an increase in hygiene quality and a decrease in food borne illnesses (Jin and Leslie, 2003, 2009; Simon et al., 2003). A similar finding was reported in Toronto, with disclosure improving compliance and decreasing foodborne illnesses (Serapiglia et al. 2007). In another study, a 'letter-grade' score on restaurant hygiene in New York showed an increase in hygiene conditions (Wong et al. 2015). The use

of a food inspection and restaurant grading system, but this time in Brazil for the 2014 FIFA world cup, also showed an improvement in food safety (da Cunha et al., 2016). A case study of Preston City Council likewise showed that cooperation improved food hygiene and safety compliance, and, in doing so, produced an economic benefit by facilitating a more efficient use of resources (Bradford-Knox et al., 2016). In Berlin, the use of a points-based scale for food safety laws again showed a positive relationship (Fietz et al., 2018). Similarly, the positive impact of public disclosure of inspection scores was also reported in a recent study of King County in Washington State, where there was a small but positive impact on inspection scores and hygiene violations (Patel and Rietveld, 2020). This is repeated in a study by Yu and Costanigro (2019), where the introduction of online restaurant disclosure scores for Orange County in North Carolina resulted in improved hygiene safety scores. Within this generally positive picture, it is important to note that in an analysis by Ho, Ashwood and Handan-Nader (2019), which looked again at the introduction of food hygiene scores in Los Angeles, the health improvements reported by Jin and Leslie (2003) were absent when considered against new data and a revised methodology. However, this, in turn, is at odds with a recent study in Northern Ireland, which provided a clear and positive association between inspection scores and foodborne illness (Fleetwood et al., 2019).

On the basis that the evidence shows improved food hygiene and safety scores reducing foodborne illness, it is apposite that we investigate how food hygiene and safety scores can be improved. A study by Yapp and Fairman (2005), which looked at compliance with food safety legislation in small businesses, found that regulators and small business viewed compliance in different ways: regulators see compliance as an ongoing proactive process of evaluation and monitoring; while small businesses see compliance as a reactive process that

follows regulator visits and inspections. This indicates that the regulator, as a component in the complex set of factors that affect compliance, can affect those factors via the regulator-regulatee interaction. And if the regulator uses education, persuasion and cooperation to achieve compliance in the first instance, which is widely held to be the case (Bardach and Kagan, 1982; Braithwaite, 1985; Cranston, 1979; Hawkins, 1984, 2002; Hawkins and Hutter, 1993; Hutter, 1988, 1989, 1997; May, 2005; May and Wood, 2003; May and Winter, 2011; Pautz, 2009, 2010), it is appropriate to ask whether this is the best approach for improving food hygiene and safety scores. Indeed, even if education, persuasion, and cooperation are not the best approach, and the regulator-regulatee interaction is a much more nuanced interaction, it is incumbent on us to ask how the interaction affects food hygiene and safety scores – to see whether we can improve the scoring to reduce foodborne illness. In the present study, even though the four neighbouring LAs under investigation collaborate and implement the same national regulations, there is a marked difference in approach to ensuring compliance. This difference is explained in the next section.

3. Four LAs: two distinctive approaches to ensuring compliance

In the UK, local authorities have some discretion over how to implement national legislation, and accordingly can develop distinctive approaches to enforcement. Indeed, the scope of this study was determined by a joint decision by the four local authorities to compare the efficacy of two distinctly different approaches that had arisen within their number: in LA1, LA2 and LA3 the approaches were broadly the same by being *compliance-to-support* orientated, with an assessment of compliance based on the conditions found during the inspection; whereas in LA4 a *support-to-compliance* approach was used, involving a number of visits to provide support and advice followed by an assessment of compliance once the establishment had had

the opportunity to implement the guidance given. (These characterizations of the respective approaches were suggested by the authors of the present study and were endorsed, along with the following explanations, by the practitioner group – representing all four LAs – that commissioned the research reported in this article.)

The rating process, as specified in the Food Law Code of Practice (England) 2017 (the Code), commences once the establishment has registered with the local authority, which they must do 28 days prior to commencement of trading. This is followed by an unannounced on-site visit to verify that systems, procedures, and documentation are in place. In LA1, LA2 and LA3 a rating would then be given followed by advice and support. Subsequent visits, which are unannounced and take place at frequencies determined by the level of risk involved, then re-inspect and re-rate. When remedial action is required and the business response is inadequate, local authorities have the power to apply written warnings and potential enforcement actions including seizure and detention of food, hygiene improvement notices, hygiene emergency prohibition notices, cautions, and prosecutions. We refer to this as a *compliance-to-support* approach, since advice and support towards compliance or an improved rating score are provided *after* the assessment of compliance is undertaken and recorded. However, a different approach is followed in LA4. It moved away from *compliance-to-support* as this is perceived, in line with the criticism of deterrence theory, to intimidate establishments, conditioning an evasive, non-cooperative approach and an often-unequal relationship.

In LA4, these considerations resulted in an alternative *support-to-compliance* approach. This was considered to comply with statutory requirements (i.e., the rating is eventually applied

following an unannounced visit, and enforcement powers are used when needed) but gives greater weight to *prior* observation of practices, identification of risks and provision of support in the compliance process. In practice, the support-to-compliance approach involves the pre-booking of visits to ensure that relevant people would be present, the use of chefs' overalls rather than white coats, and, as part of a less formal approach, the development of a more equal relationship based on trust, with an emphasis on joint working to identify solutions to improve practices. This 'hands on' approach takes place at the establishment's premises, with reports being jointly produced and agreed on-site, and actions likewise being immediately prioritized by both parties. Follow up visits continue until the establishment is deemed 'safe' and ready for a rating. Importantly, although this may require more visits and more time spent with establishments, the authority felt that it provided a more reliable indicator of food hygiene compliance, since an officer will have had the opportunity to observe a range of activities and food handlers, and to have had discussions with the Food Business Officer (FBO), and will therefore be able to make a more informed judgement in the key areas pertaining to food hygiene ('Procedures', 'Structure' and 'Confidence in Management'). (The FBO is defined in Article 3 of Regulation (EC) No 178/2002 as the natural or legal person responsible for ensuring that the requirements of food law are met within the food business under their control.) The benefit is said to be a more accurate judgement of the risks involved, which is perceived to be essential when the Code necessitates the determination of a minimum inspection interval based on the risk rating given by the inspecting officer. Hence, the LA4 approach, in preference to what is seen as a less accurate 'snapshot' from the inspection checklist approach advocated in the Code (which may or may not involve the FBO being present), considers that it is introducing positive changes in food hygiene practices and, as such, that positive changes will more likely be sustained over time.

These approaches have evolved over time, since policy changes take time to develop and to take effect. The consensus among the senior practitioners who advised and gave regular feedback on the project from which the present study originated – all responsible for implementing food hygiene regulations in their respective authorities – is that the available data span covered the conception, inception, development and hence divergence of the LA4 approach from the relatively homogeneous approach of the other three authorities. Accordingly, we cannot identify a clear division between some period before and some period after the introduction of the LA4 approach. Therefore, to investigate whether the LA4 approach gives rise to more compliant establishments, we compare the effects on establishment compliance of location in each authority (i) for each year from 2009-10 to 2017-18 relative to the baseline year 2008-09 and (ii) for each year from 2013-14 to 2017-18 relative to the five-year baseline period 2008-09 to 2012-13. However, we then combine our measure of compliance with a measure of resource expended to create a joint measure of 'compliance return', which we propose as a more comprehensive measure of the efficacy of different regulator-regulatee interactions.

4. Methodology: Research Questions; the 'Compliance return' Construct; and Data

In this Section, we set out the components of our 'compliance return' construct – i.e., the compliance level and the resource level – together with the data used to measure each component. On this platform, we address two research questions.

4.1. Research Questions

Our measure of the compliance level of each enterprise across the four authorities and in each year of the sample period enables us to address out first Research Question (RQ1):

- '*How does each authority differ in terms of compliance?*' This will identify systematic differences in compliance level (if any) between the '*compliance-to-support*' approach of LA1, LA2, and LA3 and the '*support to compliance*' approach of LA4.

Next, we measure the resource level available to each authority in each year. We then combine our measure of the resource level with our measure of the compliance level to create a joint measure of 'compliance return'. This enables us to address our second Research Question (RQ2):

- *'How are LA differences in resources expended related to compliance?'* This will identify the terms of the trade-off (if any) between compliance levels and resource levels.

4.2. The construct

We propose the 'compliance return' construct to measure jointly (i) enterprise compliance (food hygiene and safety scores) and (ii) the resource inputs made available by each local authorities to administer the system. Compared to a poorer performing 'compliance return', an improved 'compliance return' can be achieved by either an improved compliance score or less administrative input.

The administrative cost to the regulator includes the level of monitoring required to verify the compliance of the business (e.g., number of inspections, time and resources needed to investigate compliance) and the cost of enforcing non-compliance (e.g., improvement notifications, prosecution expenditure). We anticipate that deterrence would necessitate a stronger emphasis on detecting non-compliance and the collection of evidence to prove it, and the time and expense when prosecution is pursued, while cooperation would require the time and resources to create the conditions needed to avoid non-conformance in the first place (Oded, 2013; Rechtschaffen, 1998). Given the extent of these resource needs, it is often presumed, for the level of compliance being pursued, that the regulator has the necessary cognitive, computational, and administrative resources to secure compliance (Garrie and Keller, 1993; Glachant et al., 2013; Russel et al., 1986). Yet, challenging the assumption that regulators have the resources to regulate, UK regulatory authorities, in line with other countries, have experienced an ongoing downward pressure on their budgets (Regulating Futures Review, 2017). The resources of the regulator, we posit, are therefore critical to compliance no matter how the regulation is implemented. Moreover, the task is operationalized through street-level inspectors, as they are the main mode for delivering policy (Hill and Hupe, 2002; Lipsky, 1980; Meyers and Vorsanger, 2003). This also applies to compliance with food hygiene and safety standards, since this too depends on "street-level" inspections. Hence, in all four local authorities under investigation, food hygiene and safety officers are the key unit of resource when comparing the resource requirements of different approaches to regulation.

Together, measures of compliance level (i.e., the food hygiene and safety score) and of the resource level (i.e., number of food hygiene and safety officers) provide a feasible and low-

cost empirical platform for determining the overall ‘compliance return’, because both can be clearly defined and are readily available from a data acquisition perspective. We explain below, in Section 5.2, the technical procedure for combining the different units of measurement of compliance and resources into a single composite ‘compliance return’ measure.

4.3. Data

The available data enables us to measure both compliance outcomes and the corresponding resource inputs.

4.3.1. Compliance outcomes

The data on food hygiene compliance comes from inspections by professionally qualified food safety officers. Inspections take place in accordance with national statutes and thus give rise to standardized data in electronic form. The data for all four authorities is compiled and curated in a common format by the IT Department of LA3, who made the data available to the authors of the present study.

The data available enabled us to investigate the food hygiene and safety rating scheme, as indicated, over a 10-year period (fiscal years 2008/09 – 2017/18) and incorporated all inspected establishments from all four local authorities. Nationally, most establishments (90.2 percent) are rated at 3 or above, showing that compliance is generally satisfactory or better (Food Standards Agency 2018). This is also a characteristic of our sample (89.7 percent).

To assess levels and changes in food hygiene and safety compliance, we use the 'Overall Risk' score as the unit of analysis. The Overall Risk score is the total numerical score allocated to a food business establishment, following an inspection, and captures the factors that can be affected by the establishment itself: i.e., food hygiene and safety procedures; structure; and confidence in management. The Overall Risk score is continuous from 0 to 147, with the score decreasing as compliance improves (so a score of 147 is the worst possible score while 0 is the best possible score), and translates into the 0 to 5 public food hygiene rating described above. (In our sample [N= 26,285], per rating category [0 to 5] the mean Overall Risk scores [0 to 147, rounded] are: 0 – 97 [n=114]; 1 – 77 [n=1,587]; 2 – 64 [n=1,004]; 3 – 56 [n=5,181]; 4 – 47 [n=4,753]; and 5 – 37 [n=13,646]. The correlation is near perfect (-0.9965), noting that the higher the better for the publicly available rating scale but the lower the better for the underlying Risk Score.) The Overall risk score is thus the measure with the highest informational content on the food hygiene and safety performance of the establishment, and therefore is our preferred measure of the food hygiene and safety compliance of the establishment.

A concern is that scoring differences between the four local authorities under consideration may reflect systematic differences in assessment and recording practices, even in the context of national regulations. Indeed, the UK Food Standards Agency (2015) has investigated just this issue, concluding that "management practices within local authorities may result in inconsistencies observed within the case study pairs". However, our study is not based on statistically matched but otherwise unrelated "study pairs", but on a group of local authorities in close geographical proximity that work closely together and share infrastructure (such as the IT system that enabled the data for the present paper to be accessed in a form suitable for

econometric analysis). Indeed, the report also concludes that "where present, the use of shared processes and policies to enforce food safety regulation has promoted regional consistency in the delivery of official controls between local authorities". This characterizes the relationship between the four local authorities under investigation; indeed, the project from which the present study originates is an outcome of this relationship. Close collaboration of the four LAs under investigation together with the audit trail outlined above suggests common and accurate recording practices and, hence, the validity of our compliance measure.

Our data on the Overall Risk score is complete for all four LAs throughout the sample period, covering all inspections in each local authority over a 10-year period, from fiscal year 2008-09 to 2017-18. In total, there are 11,294 establishments uniquely identified over time. The following are the percentages of the total number of inspections accounted for by the different categories of establishments subject to inspection: Distributors/Transporters (1.45 percent); Slaughterhouse (0.11 percent); Importers/Exporters (2.47 percent); Manufacturers and Packers (2.16 percent); Primary producers (0.21 percent); Restaurants and Caterers (72.74 percent); Retailers (21.42 percent); and Supermarket/Hypermarket (1.96 percent). The number of establishments by local authority was as follows: 948 in LA4; 1,786 in LA3; 1,591 in LA2; and 6,969 in LA1. For the analysis of food hygiene and safety compliance, this involved 26,285 observations from individual inspections, with a maximum of six observations for any one establishment and a mean of 2.3 observations per establishment across all 11,294 establishments. As Appendix B shows, the observations were spread fairly evenly across local authorities and by year.

4.3.2. Resource inputs

LA3 also provided data from the respective personnel departments of the four LAs. This enabled the creation of a new variable to measure resource inputs: *Resource* – defined as the ratio of the number of available inspectors (full-time equivalents, henceforth FTE) to the number of inspections carried out by each local authority in each year (= FTE / Number of inspections).

Descriptive data by authority and year on the number of inspections and the available FTE are reported in Tables 1 and 2. The derived *Resource* variable is reported in Appendix C.

	2008-09	**2009-10**	**2010-11**	**2011-12**	**2012-13**	**2013-14**	**2014-15**	**2015-16**	**2016-17**	**2017-18**
LA4	3.05	2.25	1.86	3	3	3	3	3	3	3
LA3	4.5	4	4	4	4	4	4	4	4	4
LA2	3.1	2.81	3.72	3.4	3.4	3.4	3.4	3.4	2.5	2.5
LA1	7.5	7.5	7.5	7	6	5	5	4.6	4.6	4.6

Table 1: FTE numbers by Authority and Year

	2008-09	**2009-10**	**2010-11**	**2011-12**	**2012-13**	**2013-14**	**2014-15**	**2015-16**	**2016-17**	**2017-18**
LA4	90	355	265	334	426	248	184	501	311	372
LA3	204	173	237	1,277	560	552	541	602	508	485
LA2	219	224	370	547	1,139	668	673	691	663	819
LA1	657	971	856	839	666	831	690	820	847	4,870

Table 2: Number of inspections by Authority and Year

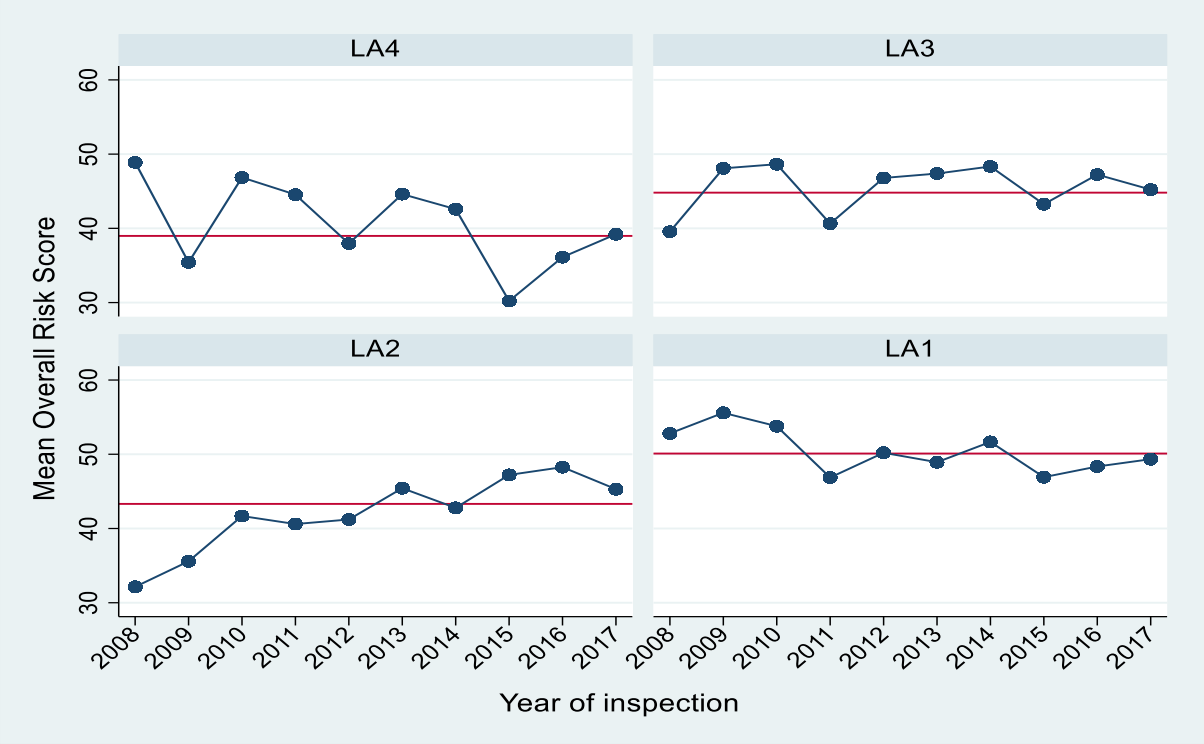
Salary data were not available throughout the sample period. However, standard professional qualifications and salary scales support the interpretation of our *Resource* variable as a proxy for cost.

5. 'Compliance return': Empirical application

First, we use econometric estimation to identify the compliance effects of each LA in each year relative to both a baseline year and a baseline period. Second, we combine our measure of compliance with our measure of resource input to obtain our empirical 'compliance return' construct, which is the compliance output return per unit of resource input. Finally, we estimate the effects of each LA's approach to implementation on 'compliance return'.

5.1. Estimating Relative Compliance Effects: Model and Results

We gain some initial insight into the effects of the different approaches to compliance by comparing the mean Overall Risk scores for each local authority over the whole sample. The mean Overall Risk Score for all 26,285 observations is 46.21. Figure 1 displays the mean Overall Risk Score from all inspections undertaken by each local authority in each fiscal year (2008-09 to 2017-18). The horizontal red lines mark the mean for each authority over the sample period.



Note: Horizontal red lines indicate the mean for each local authority over the sample period.

Figure 1. Mean Overall Risk Score by year (2008-09 to 2017-18) and local authority

The authority means over the whole sample period vary with high degrees of statistical significance: LA1=50.10 > LA3=44.81 > LA2=43.32 > LA4=38.98 (t-tests reject the null of equal means in each comparison with p=0.0000). These unconditional (i.e., model-free) statistics suggest that establishments in LA4 typically display a higher level of compliance than do establishments in the other three LAs, and are thus consistent with the suggestion that LA4's *support-to-compliance* approach may outperform the *compliance-to-support* approach of the other authorities.

Next, we specify an econometric model to isolate and measure (i.e., to identify) the influence on establishment compliance of location in a particular authority in each year separately

relative to the base year, 2008-09. The estimated establishment compliance model (Equation 1) explains variation in Overall Risk scores (the dependent variable) by variations in the independent variables, where i = 1, …, 11,294 indexes the establishments and t = 2009-2010, …, 2017-18 indexes the years in which inspections occurred (2008-09 is omitted as the base year).

$$Overall\ Risk\ Score_{it} = \widehat{Constant} + \sum_{t=2009-10}^{2017-18} \hat{\beta}_{1t} DVLA1 + \sum_{t=2009-10}^{2017-18} \hat{\beta}_{2t} DVLA2 + \sum_{t=2009-10}^{2017-18} \hat{\beta}_{3t} DVLA3 + \sum_{t=2009-10}^{2017-18} \hat{\beta}_{4t} DVLA4 + \hat{\beta}_5 Restaurants\ \&\ Caterers_i + \hat{\alpha}_i + \hat{\varepsilon}_{it} \quad (1)$$

DVLA1 is a dummy variable (DV) (=1 for each inspection by LA1; =0 for each inspection by another local authority) and $\sum_{t=2009-10}^{2017-18} \beta_{1t} DVLA1$ sums the estimated effects ($\hat{\beta}_{1t}$) – relative to the base year, 2008-09 – in each year of an establishment being subject to regulation by LA1 – i.e. $\hat{\beta}_{1,2009\text{-}10}$DVLA1 + $\hat{\beta}_{1,2010\text{-}11}$DVLA1 + … + $\hat{\beta}_{1,2017\text{-}18}$DVLA1. By analogy, $\hat{\beta}_{2t}$, $\hat{\beta}_{3t}$ and $\hat{\beta}_{4t}$ respectively estimate the effects for each year of an establishment being subject to regulation by LAs 2, 3 and 4. Only one establishment category varies over time in our sample and can thus be included in the model (because, otherwise, lack of time variation gives rise to perfect collinearity in the data): accordingly, we control for whether the establishment belongs to the "Restaurants and Caterers" category (=1; otherwise, 0). The model also estimates a "fixed effect" for each establishment ($\hat{\alpha}_i$), which controls for all constant (or slowly moving) but unobserved between-establishment influences on compliance. Each establishment fixed effect not only controls for unique features of the establishment itself but also for (i) higher-level influences on the establishment – such as the socio-economic environment in a local authority area – that affect all establishments in a particular local authority in a similar way throughout the sample period, and (ii) initial conditions for establishments, local authorities and local authority areas at the onset of, and

thus constant throughout, the sample period. However, these fixed effects do not account for the effect of the Local Authority by which the establishment is regulated, because our variables of interest allow this influence to be time varying. Accordingly, whereas the fixed effects account for all the variation *between* establishments, the estimated effects of our variables of interest – the authority-year dummies – arise entirely from changes *within* establishments over time. The "Constant" is the mean of the fixed effects, thereby capturing any systematic unobserved influences that affect all establishments in all periods. Finally, the usual regression errors ($\hat{\varepsilon}_{it}$) capture unobservable idiosyncratic influences on each establishment in each period. The complete estimated model is reported and interpreted in detail in Appendix A. Here, we focus on the estimates of interest, i.e. $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$, $\hat{\beta}_{3t}$ and $\hat{\beta}_{4t}$. Respectively, for LA1, LA2, LA3 and LA4, these estimate the year-by-year changes in average Overall Risk score – measuring establishment compliance – compared to the average Overall Risk score in 2008-09 (the base year). Figure 2a displays these estimates graphically.

In most years, all four authorities record improved compliance. The reference line at zero indicates no change; negative values on the Y-axis show improvements; and, conversely, positive values on the Y-axis show decreases in average compliance. The dots represent yearly changes in average overall risk score relative to the base year of 2008-09, while the vertical bars running from the dots are 95 percent confidence intervals, which show the range within which we can be 95 percent sure that that the improvement is valid (i.e., not just a random fluctuation). The confidence intervals are calculated from the cluster-robust standard errors reported in Appendix A, Table A1. (Likewise for the results reported in Figures 2, 3, and 4 below.) When the confidence intervals touch or cross the zero reference line then the estimate is not regarded as statistically significant (i.e., not distinguishable from zero).

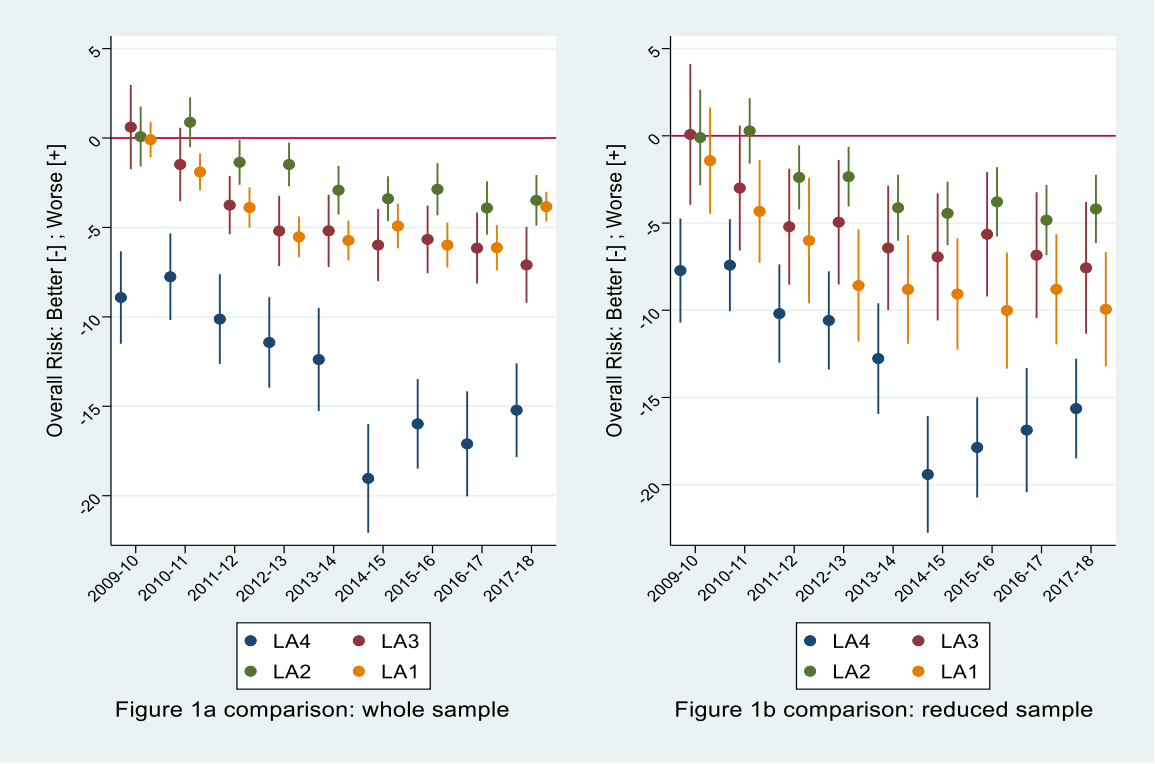


Figure 1a comparison: whole sample

Figure 1b comparison: reduced sample

Source: Estimated fixed effects models. The Panel 2a model is reported and explained in full in Appendix A.

Figures 2a and 2b. Food Hygiene and Safety Compliance by Local Authority – Changes in average Overall Risk score from 2008-09 to 2017-18 (each year's estimate is relative to the base year of 2008-09):

- **2a – full sample (11,294 establishments; 26,285 observations);**
- **2b – establishments with 5 or 6 inspections (1,776 establishments; 8,951 observations)**

Each dot depicts one of the respective estimated coefficients $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$, $\hat{\beta}_{3t}$ and $\hat{\beta}_{4t}$ (i.e., one for each local authority in each year), while the vertical bars depict the associated confidence intervals (such that the shorter the bar the more precise the estimate). In general, the greater the number of observations, the more precise are statistical estimates. Accordingly, the confidence intervals for LA4 extend further than do those of the other local authorities, because there are fewer establishments and observations available for LA4 (948 and 3,086)

than for the other local authorities: LA2 (1,591 and 6,013); LA3 (1,786 and 5,139); and LA1 (6,969 and 12,047).

The Authority-Year effects reported in Figure 2a are not cumulative and can be scaled against the mean Overall Risk score (= 46.21).

- LA4 improved establishment compliance in all 9 years: in each year the improvement on the Overall Risk score was substantial, being between -8.92 and -19.03 on the base year (i.e., between -19.3 percent and -41.2 percent of the mean Overall Risk score).
- LA3 improved establishment compliance in 7 of the 9 years, with improvements between -3.75 and -7.10 on the base year (i.e., between -8.1 percent and -15.4 percent of the mean).
- LA2 improved establishment compliance in 7 of the 9 years, although the improvement was less pronounced, being between -1.35 and -3.39 on the base year (i.e., between -3.0 percent and -8.5 percent of the mean).
- LA1 improved establishment compliance in 8 of the 9 years, being between -1.90 and -6.13 on the base year (i.e., between -4.1 percent and -13.3 percent of the mean).

From these estimates, we conclude firstly that not only is the average level of compliance higher in LA4 than in the other three authorities (Figure 1) but that this effect is systematic, present in every year. While all four local authorities improved establishment compliance, LA4 achieved, in each year and by some margin, a greater improvement relative to the base year than did each of the other three local authorities. Whereas the upper bounds of the LA4 95 per cent confidence bands are clearly separated (and thus statistically distinct) from the

lower bounds for the other local authorities in every year, the mainly overlapping confidence intervals between the LA1, LA2 and LA3 estimates show that with respect to improvements in compliance these three authorities have more in common with each other than with LA4.

Together with the superior average compliance in LA4 (Figure 1), the persistence of greater compliance throughout the sample period (Figure 2) is evidence consistent with the greater effectiveness of LA4's *support-to-compliance* approach in comparison with the *compliance-to-support* approach used by the other authorities. However, best practice considerations suggest that we should investigate the possibility that features of our data or model might be accounting for the effects – and their trajectories – presented in Figure 2a. Accordingly, we investigate the robustness of these findings with respect to two potential sources of bias.

1. The widely varying number of inspections per establishment could be a source of bias.
2. Time-series comparisons – especially over short periods – are notoriously sensitive to initial conditions. Figure 1 shows a markedly worse mean Overall Risk Score for LA4 at the beginning of the sample period, which raises the possibility that subsequently superior compliance reflects only a poor starting position.

In the whole sample used so far, the average number of inspections is 2.3, ranging between one and six (see Appendix A). For comparison, we estimate our model on a reduced sample, comprising only those establishments with five or six inspections, respectively accounting for 8,525 and 426 observations (together a little over one-third of the whole sample). For individual local authorities, the percentage reductions in the number of establishments and observations available for estimation are: LA1 – by 96 percent and 89 percent, respectively;

LA2 – 52 percent and 36 percent; LA3 – 76 percent and 57 percent; and LA4 by 68 percent and 49 percent. Comparison of Figure 2b (reduced sample) with 2a (full sample) shows that, while not identical, the overall pattern of estimates is robust to large changes in the number of establishments and observations. First, in each year, the point estimates for LA4 continue to be much the same size and lower than for the other three authorities. The upper bounds of the LA4 95 per cent confidence bands are clearly separated (and thus statistically distinct) from the lower bounds for the other LAs in four from nine years (and borderline in another) compared to lower in every year. This less clear separation of LA4 from the other three authorities reflects (i) lower estimates for LA1, conditioned by the most extreme reduction in sample size, and (ii) a consequently smaller sample resulting in less precise estimates (i.e. wider confidence bands). (In the reduced sample, the number of establishments and observations for each authority are: LA1 – 270 and 1,350; LA2 – 762 and 3,841; LA3 – 436 and 2,195; and LA4 – 308 and 1,565.) Nonetheless, while the LA1 estimates are lower in the reduced sample, overlapping confidence intervals between the LA1, LA2 and LA3 estimates continue to show that with respect to improvements in compliance these three authorities have more in common with each other than with LA4. Accordingly, we continue our analysis using the whole sample.

To check robustness with respect to the baseline, we estimate our model on the full sample. However, instead of estimating our authority-year variables relative to a single year (2008-09) we estimate our variables of interest relative to the five-year period from 2008-09 to 2012-13. The full model is available on request. The estimates of interest are displayed graphically in Figure 3.

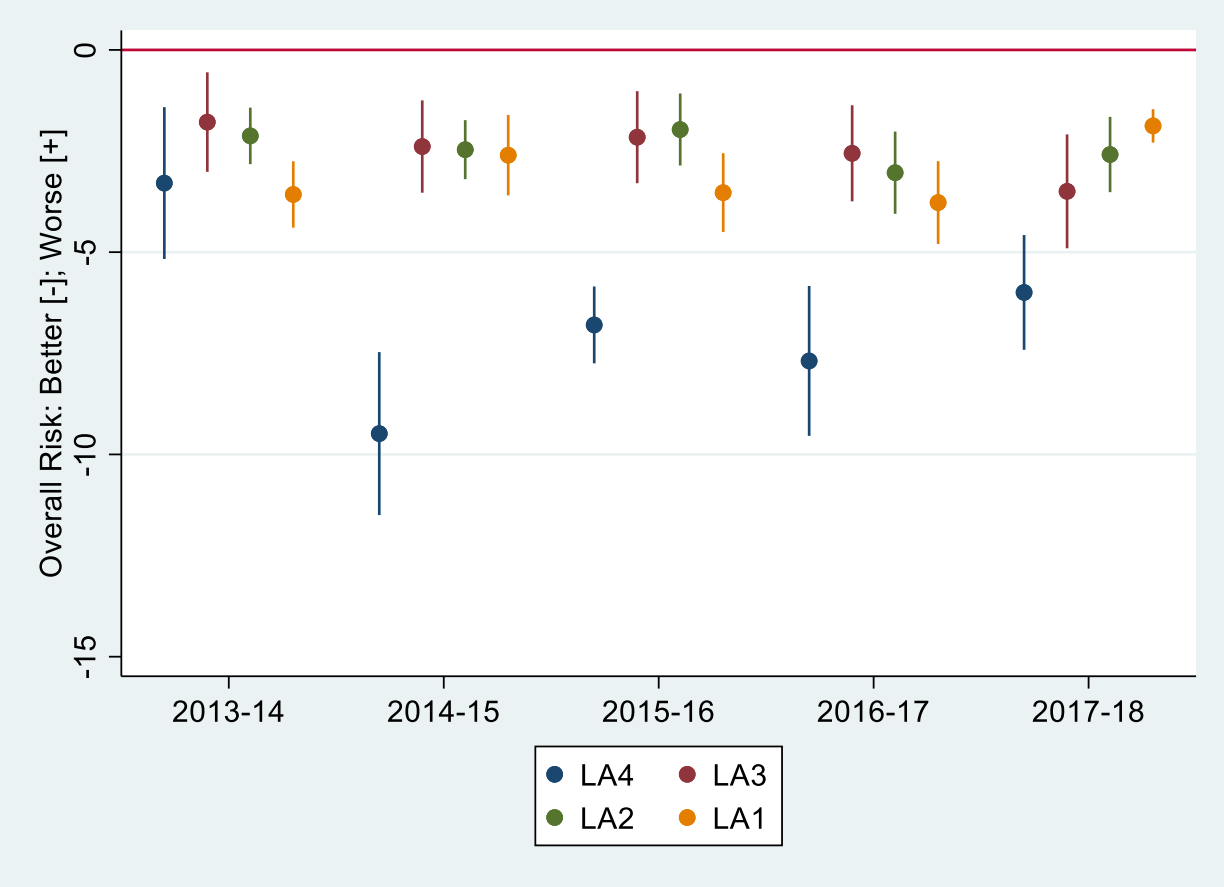


Figure 3. Food Hygiene and Safety Compliance by Local Authority – Changes in average Overall Risk score from 2008-09 to 2017-18 (each year's estimate is relative to the base period from 2008-09 to 2012-13)

Qualitatively, the results are similar to those estimated relative to the single baseline year, 2008-09: in each year, the LA1, LA2 and LA3 estimates are not statistically different from each other, whereas four from five LA4 estimates show statistically greater improvements (three with clear and one with borderline statistical significance). Quantitatively, however, the range of improvements is reduced from between -8.92 percent and -19.03 on the base year 2008-09 to between -3.29 percent and -9.48 percent (i.e., between -7.2 percent and -20.5 percent of the sample mean). If we extend the base period to the first six years, the improvements range between -5.44 percent and -9.01 percent. In round terms, diluting the

effect of early years with higher (worse) Overall Risk Scores reduces LA4's outperformance by about a half.

5.2. Resource input and the measurement of compliance return

So far, we have investigated variations in compliance associated with different approaches followed by the four local authorities. We have presented evidence (i) that LA4's '*support-to-compliance*' approach is more effective in achieving compliance than the '*compliance-to-support*' approach used by the other authorities, and (ii) that this outperformance is not an artifact of data limitations or the choice of baseline. In this Section, we investigate the influence of differing levels of resource input on differing levels of compliance achieved by the four LAs.

This part of the investigation proceeds in two stages: first, we construct our measure of 'compliance return' to capture the combined effect of *Overall Risk Score* (our compliance output variable) and *Resource* (our resource input variable); second, we use this joint measure as a new dependent variable in our econometric model (Eq.1).

Stage 1.

- Step 1. The values of our compliance variable – *Overall Risk Score* – are transformed into z scores; *z_Overall Risk*.
- Step 2. The values of our *Resource* variable (see above: Section 4.3.2) are likewise transformed into z scores; z_*Resource*.

- Step 3. *Overall Risk Score* and *Resource* are measured in different units. However, once standardized by z transformation, we can follow Chignell et al. (2015) to create a composite performance measure by summation of the respective z scores:

$$Compliance\ Return_{it} = 0.5\ .\ z\_Overall\ Risk_{it} + 0.5\ .\ z\_Resource_{at} \qquad (2)$$

Subscripts *i* and *t* – as in Equation 1 – indicate that *Compliance Return* and z*_Overall Risk* vary by establishment and year, whereas z*_Resource* varies by local authority (subscript *a* = LA1, LA2, LA3 and LA4) and year. *Compliance Return* is interpreted in the same manner as *z_Overall Risk*, which enters with a positive sign so that higher (lower) values are registered as worse (better) *Compliance Return*. In Section 4.3.2, we defined our *Resource* variable as FTE / Number of inspections. Hence, for consistency, *z_Resource* is also entered positively in Eq.2, as (i) more FTE per inspection means greater resource for a given *Overall Risk* score, so reduced efficiency and a higher (worse) *Compliance Return* score, while (ii) more inspections for a given FTE count increases efficiency and thus lowers (improves) the *Compliance Return* score. Weighting allows greater or less influence to be attributed to the two components. However, departing from equally weighted components (both 0.5) must be guided by "a comprehensive set of research findings" (Chignell et al., 2015: 38), which do not exist in the present case and go beyond the scope of the present study.

- Step 4. *Compliance Return* is also transformed into a z variable, according to the following formula. (The method is clearly set out at: https://stats.stackexchange.com/questions/348192/combining-z-scores-by-weighted-average-sanity-check-please.) Equation 3 is an approximation, because it assumes that the component z scores are independent. However, the correlation is very small (-0.0230) and so makes no practical difference to the calculation. The resulting composite measure is

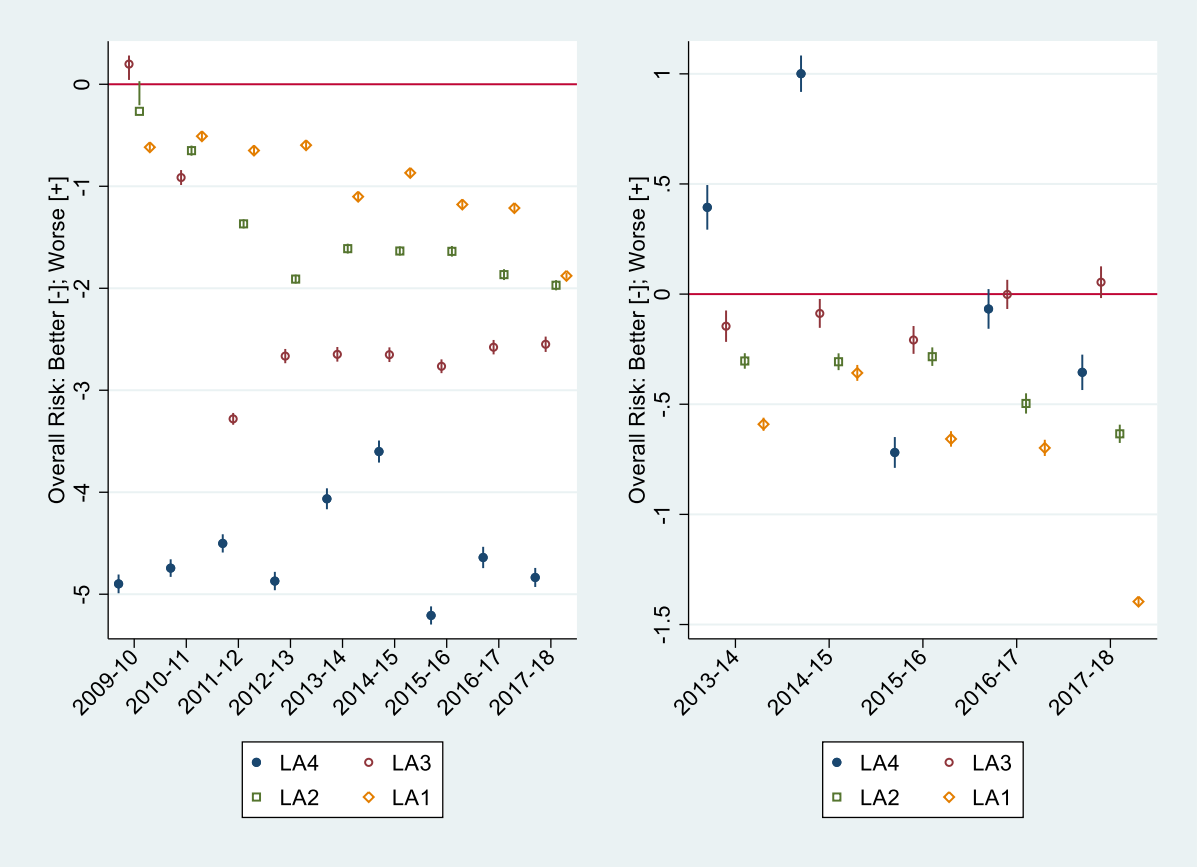
standard normal distributed with mean zero (-5.93e-10) and standard deviation very close to one (0.988).

$$Z_W = \frac{\frac{w_A}{w_A + w_B} Z_A + \frac{w_B}{w_A + w_B} Z_B}{\sqrt{\left(\frac{w_A}{w_A + w_B}\right)^2 + \left(\frac{w_B}{w_A + w_B}\right)^2}} \qquad (3)$$

where Z_W is our standardized *Compliance Return* variable (z_*Compliance Return*), a weighted combination of Z_A (standardized *Overall Risk* score, z_*Overall Risk*), weighted w_A, and Z_B (standardized *Resource*, z_*Resource*), weighted w_B (where in our case $w_A = w_B = 0.5$).

Stage 2.

The standardized *Compliance Return* variable (Z_W or z_*Compliance Return*) is then used as the dependent variable in the model (Eq.1), which otherwise has the same specification as our model reported in Appendix A and informing Figures 2 (a and b) and 3. The estimated effects on our standardized z_*Compliance Return* variable of regulation by each respective authority in each year are set out in Figure 4: with respect to the base year, in Figure 4a; and with respect to the five-year base period, in Figure 4b. (The estimated models are available on request.)



Note. Bars within the hollow markers are confidence intervals, which reflect very precise estimation.

Figure 4. Local Authority Food Hygiene and Safety: Standardized Composite *Compliance Return* (Standardized *Overall Risk* score adjusted for Standardized *Resources*)

- **Panel a: from 2009-10 to 2017-18 (relative to 2008-09)**
- **Panel b: from 2013-14 to 2017-18 (relative to 2008-09 – 2012-13)**

The message of Figure 4 is that the systematically greater compliance achieved by LA4 (see Figures 2 and 3) comes at a cost. Although Figure 4a indicates superior performance from LA4, this reflects a very high z_*Compliance Return* value in the base year (see Appendix C, Table C1), which, in turn, is driven by a very low number of inspections in 2008-09 (see Table 2). Accordingly, we regard Figure 4b as presenting the more reliable estimates, as these

measure performance effects relative to a five-year base period in which the disproportionate effect from 2008-09 is substantially diluted.

6. Discussion

In Equation 3, our procedure of applying equal weights to compliance (z_*Overall Risk*) and resources (z_*Resource*) in calculating our composite performance measure (z_*Compliance Return*) is sufficient to transform the typically greater establishment compliance in LA4 (Figures 2 and 3) into 'compliance return' measures – accounting for both compliance output and resource input – that are either no better or substantially worse than those estimated for the other authorities (Figure 4b). Conversely, the performance estimates now flatter LA1, which typically had a lower-than-average ratio of inspectors (FTE) to the number of inspections, although the apparently outlying performance in 2017-18 is driven by an outlying number of inspections in that year (Tables 1 and 2 and Appendix C). Of course, different comparisons could be obtained by imposing different weights: weighting Resource at zero gives a 'compliance return' measure reflecting only compliance (the *Overall Risk* score reported in Figures 2a, 2b and 3), which suggests outperformance by LA4's *support to compliance* approach; conversely, increasing the weight of *Resource* yields 'compliance return*'* measures that increasingly temper the benefits from compliance by taking into account the resource implications (Figure 4), which casts the *compliance to support* approach of the other three LAs in a more favorable light. Compared to LA1, LA2 and LA3, LA4 was able to significantly improve establishment compliance using a *support-to-compliance* approach but only on the basis of a stable FTE that was relatively high in relation to the number of inspections (Tables 1 and 2, and Appendix C). Conversely, the *compliance-to-support* approach combines lower levels of improved compliance with lower costs.

In effect, we have introduced an apparatus for exploring the implications of different assumptions regarding, for example, the respective claims of the public health benefits of increased compliance and the costs of delivery. However, the theoretical and empirical – not to mention political – justification of different weighting schemes is beyond the scope of this paper. Here, it is sufficient to note that, once we look past the distorting influence of the base year, these estimates confirm that the benefits of greater compliance from the LA4 approach come at the cost of greater resources.

7. Policy Implications and Conclusions

This study is designed to support practitioner debate and decision making by evaluating the associations between implementation approaches, establishment compliance and resource implications. To this end, we analysed a large set of data on food hygiene and safety inspections carried out by four UK local authorities in close geographical proximity, over a 10-year period, to estimate the relative effects of different approaches to implementation on the 'compliance return'. We introduced the construct of 'compliance return' to embrace two elements: (i) compliance effects; and (ii) resource implications. Evaluation of the relative efficacy of different approaches to regulation through the lens of 'compliance return' does not provide a comprehensive cost-benefit analysis but is both more comprehensive than a focus on either compliance or resources in isolation and – above all – feasible within the limitations of the available data. Data on compliance and corresponding resource requirements are available internally, although, of course, measured in different units. On this platform, we advance a methodology to combine these components into a single

compound measure of 'compliance return', which is a construct or tool for (i) gaining insight into the tradeoff between compliance and cost and (ii) taking account of this tradeoff in evaluating regulatory regimes where there is a choice between more resource-intensive cooperative approaches and less resource-intensive approaches emphasizing deterrence. As such, this low-cost approach to evaluation could be applied elsewhere by practitioners in food safety regulation and, possibly, in other regulated activities.

The evidence from our empirical investigation addresses our two research questions. RQ1 asked: '*How does each authority differ in terms of compliance?*' Fixed effects econometric analysis of our compliance variable (Overall Risk score) indicated that while all four local authorities were able to improve food hygiene and safety compliance at establishments in their district, LA4 stimulated the most improvement. This suggests that with respect to compliance considered in isolation (measured by inspectors' Overall Risk scores), *support-to-compliance* is more efficacious than *compliance-to-support*. Neither of these of these approaches is at the extremes of the continuum from deterrence to cooperation but, rather, are variant hybrid forms. So, in this case at least, the evidence suggests that even moderate changes in the approach to regulation can have substantial effects on compliance outcomes. Yet, in forming judgements regarding the respective efficacy of the four LAs, and, in particular, the respective merits of *support-to-compliance* and *compliance-to-support*, we need to consider RQ2: *'How are LA differences in resources expended related to compliance?'* RQ2 is addressed by translating our 'compliance return' construct into an empirical measure to take account of both compliance outcomes and resource inputs. This 'compliance return' measure is then used as the dependent variable in our econometric model to identify the effects of each local authority in each year of the sample. Focusing on the

alternative approaches, *support-to-compliance* requires more FTE resource per inspection than does *compliance-to-support*. Econometric analysis enables the terms of this tradeoff to be identified across regulatory authorities and over time. The estimates reported in the present study show that down weighting the importance of resource inputs favours the *support-to-compliance* approach (greater emphasis on cooperation), while increasing the weighting of resource inputs can present the *compliance-to-support* approach (greater emphasis on deterrence) in an increasingly favorable light. Of course, once armed with this information the choice for policy makers between different regulatory approaches is politically delicate, since cost constraints must be balanced against the health of the population.

The evaluation methodology advanced in this article may be useful whenever the benefits of compliance must be considered within the constraints of cost. It is a low-cost methodology, because it is implemented using only data already available. The technical requirements are minimal, such that the template provided in this study could easily be implemented either in-house or with help from a local university. Regarding practical implementation, we recommend – as demonstrated in this article – attention to robustness checks, in particular, the need to ensure that results do not simply reflect the choice of baseline period. Finally, the methodology should also prove easy to adapt and improve, depending on local requirements and data availability.

For public policy at the LA level, our findings suggest that different approaches to implementation, occupying – in our case – hybrid positions along the continuum from deterrence to cooperation, can be judged differently depending on whether policy makers consider (i) compliance alone or (ii) compliance jointly with resource implications. The

method of investigation advanced in this study can provide quantitative information to inform choices between different approaches.

Our methodology and results can also inform key national bodies, which in the UK include the Office for Product Safety and Standards and the Local Authority Regulatory Services Excellence Framework, which stipulates that practices should be shared to inform and improve delivery. This study could also inform the 'Regulating our Future' agenda of the UK Foods Standards Agency and their plans for reforming food industry regulation (e.g., plans around a permit to trade, segmentation based on risk, and private assurance schemes to alleviate pressure on local authority resources).

We propose our 'compliance return' construct as a low-cost approach to evaluating the relative effectiveness of different approaches to implementing food hygiene regulations. Whereas full cost-benefit analysis is likely to be ruled out on grounds of cost and feasibility alike, the *Compliance Return* approach can be implemented using existing data, hence at minimal cost. The data required are inspection results and personnel details, while analysis can be completed either in-house or with a minimum of external help. However, policy makers should be aware that the use of existing data also involves limitations. In the case of the present study, the record of inspections does not include data that would, for example, (i) enable analysis of potential gender effects – among owners, managers, or employees of inspected businesses – on compliance, or (ii) take account of the economic impact on the business of securing compliance, the importance of which was highlighted by Antle (1999) and Traill and Koenig (2010). Unfortunately, the desirability of additional data to inform a richer and more complex analysis of 'compliance return' conflicts with the requirement of

feasibility. Extending food hygiene inspections to incorporate wider business intelligence data – e.g., concerning owners, managers and employees, the number of full- and part-time employees, turnover, and so on – would enable the benefit of a more rounded and complex analysis but only at the cost of more intrusive and time-consuming inspections. For policy makers wanting to evaluate the 'compliance return' of different approaches to implementation, we propose the approach outlined in this article as a practical compromise.

Regarding the applicability of our *Compliance Return* approach beyond the implementation of food hygiene regulations in a single English region, we conclude that our approach may have external validity in regulatory fields where (i) full cost-benefit analysis is not feasible, (ii) there is some discretion – and thus some variation – regarding implementation, especially at local level, and (iii) consistent records of inspection results are maintained across implementing authorities. In this article, we do not attempt to identify such regulatory fields. Nonetheless, we have attempted to provide sufficient detail about the context and implementation of our approach to enable readers to make informed judgements as to the applicability of our evaluation methodology to their own circumstances.

Acknowledgements

Financial support was provided by the Stoke-on-Trent and Staffordshire Local Enterprise Partnership and Staffordshire University. The participating local authorities also provided significant support in kind. Special thanks are owed to the members of the Steering Group, whose support made the analysis possible. However, any shortcomings are the sole responsibility of the authors.

Author contribution

DW: Conceptualization (primary role); Investigation; Methodology (secondary role); Writing - Original Draft; Writing - Review & Editing; Visualization; Supervision; Project administration; Funding acquisition.

GP: Conceptualization (secondary role); Investigation; Methodology (primary role); Formal analysis (primary role); Validation; Data Curation; Writing - Original Draft; Writing - Review & Editing; Visualization; Supervision; Project administration.

MA: Conceptualization (secondary role); Validation; Formal analysis (secondary role); Investigation; Writing - Review & Editing; Visualization; Project administration.

According to the "percent-contribution-indicated" (PCI) approach, the respective authorial contributions are – in descending order - 42.5%, 42.5%, and 15%.

Data availability statement

The dataset, in anonymized form, generated during and/or analysed during the current study is available from the corresponding author on reasonable request. Upon publication, the data will be made available online together with the Stata syntax used to estimate the models and generate the graphical displays of results reported in this article.

Conflict of interest

Authors state no conflict of interest.

References

Andreen, W.L. (2007). Motivating Enforcement: Institutional Culture and the Clean Water Act. *Pace Environmental Law Review*, Vol. 24, 67-98.

Andreoni, J., Harbaugh, W. and Vasterlund, L. (2003). The Carrot or Stick: Rewards, Punishments, and Cooperation. *American Economic Review*, Vol. 93. No. 3. 893-902.

Antle, J.M. (1999). Benefits and Costs of Food Safety Regulation. *Food Policy*, Vol. 24, No. 6, 605-623.

Ayres, I. and Braithwaite, J. (1992). *Responsive Regulation: Transcending the Deregulation Debate*. Oxford University Press: Oxford

Bardach, E. and Kagan, R.A. (1982). *Going by the Book: The Problem of Regulatory Unreasonableness*. Temple University Press: Philadelphia.

Barnes, J., Whiley, H., Ross, K. and Smith, J. (2022). *Defining Food Safety Inspections. International Journal of Environmental Research and Public Health*, Vol. 19, No. 2, 789-807.

Bates, L., Darvell, J.M. and Watson, B. (2015). Young and unaffected by road policing strategies: Using deterrence theory to explain provisional drivers' (non)compliance. *Australian & New Zealand Journal of Criminology*, Vol. 50, No. 1, 23-38.

Becker, G.S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, Vol. 76, No. 2, 169-217.

Borraz, O., Beaussier, A.L., Wesseling, M., Demeritt, D., Rothstein, H., Hermans, M., Huber, M. and Paul, R. (2022). Why Regulators Assess Risk Differently: Regulatory Style, Business Organization, and the Varied Practice of Risk-Based Food Safety Inspectors Across the EU. *Regulation & Governance*, Vol. 16, No.1, 274-292.

Bradford-Knox, R., Kane, K. and Neighbour, S. (2016). Approaches to Food Safety Hazard Control and Risk Management: A Case Study of Preston City Council's Food Safety Compliance Strategy. *International Journal of Management and Applied Research*. Vol, 3, No. 1, 14-29.

Braga, A., Weisburd, D. and Turchan, B. (2018). Focussed deterrence strategies and crime control: An updated systematic review and meta-analysis of the empirical evidence. *Criminology & Public Policy*, Vol. 17, No. 1, 205-250.

Braithwaite, J. (1985). *To Punish or Persuade*. State University of New York Press: Albany.

Buckley, J.A. (2015). Food Safety Regulation and Small Processing: A Case Study of Interactions Between Processors and Inspectors. *Food Policy*, Vol. 51, 74-82

Burby, R.J. and Patterson, R.G. (1993). Improving Compliance with State Environmental Regulations. *Journal of Policy Analysis and Management*, Vol. 12, No. 4, 753-772.

Cabinet Office (2017). *Regulatory Futures Review*. https://www.gov.uk/government/publications/regulatory-futures-review

Chen, Y., Ramamurthy, K. and Wem, Kuang-Wei (2012). Organization's Information Security Policy: Stick or Carrot Approach? *Journal of Management Information Systems*, Vol. 29, No. 3, 157-188.

Chignell, M., Tong, T., Mizobuchi, S, Delange, T., Ho, W. and Walmsley, W. (2015). Combining Multiple Measures into a Single Figure of Merit. *Procedia Computer Science* 69, 36-43. https://www.sciencedirect.com/science/article/pii/S1877050915031683

Cohen, M.A. (2000). Empirical Research on the Deterrent Effect of Environmental Monitoring and Enforcement. *Environmental Law Reporter*, Vol. 30, 10245-10252.

Cranston, R. (1979). *Regulating Business: Law and Consumer Agencies*. Macmillan: London.

da Cunha, D.T., Saccol, A.L., Tondo, E.C., de Oliveira, A., Ginani, V.C., Araújo, V.C., Lima, T.A.S., da Castro, A.K.S. and Stedefeldt, E. (2016). Inspection Scores and Grading System for Food Services in Brazil: The Results of a Food Safety Strategy to Reduce the Risk of Foodborne Diseases During the 2014 World Cup. *Frontiers in Microbiology*. Vol. 7, No. 1, Article 614. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4847479/

Downs, G.W., Rocke, D.M. and Barsoom, P.N. (1996). Is the good news about compliance good news for cooperation? *International Organization*, Vol. 50, No. 3, 379-406.

Earnhart, D.H. and Glicksman, R.L. (2015). Coercive vs. Cooperative Enforcement: Effect of Enforcement Approach on Environmental Management. *International Review of Law and Economics*, Vol. 42, 135-146.

Ehrlich I. (1975). The Deterrent Effect of Capital Punishment: A Question of Life and Death. *American Economic Review*, Vol. 65, No. 3, 397-417.

Fietz, A., Bavorová, M., Grűner, S. and Hirschaner, N. (2018). Compliance with Food Safety Laws in Germany: Food Businesses in Berlin. *Law & Policy*, Vol. 40, No. 3, 267-285.

Finnish Council of Regulatory Impact Analysis (2017). *Annual Review*. Prime Minister's Office, Finland. https://vnk.fi/documents/10616/7861578/Finnish+Council+of+Regulatory+Impact+Analysis+Annual+Review+2017/5b9b4fd4-aa89-4700-b292-fa11b7d3cc43/Finnish+Council+of+Regulatory+Impact+Analysis+Annual+Review+2017.pdf?version=1.1

Fleetwood, J., Rahman, S., Holland, D., Millson, D., Thompson, L. and Poppy, G. (2019). As Clean as They Look? Food Hygiene Inspection Scores, Microbiological Contamination, and Foodborne Illness. *Food Control*, Vol. 96, 76-86.

Food Law Code of Practice (England) (2017). https://www.food.gov.uk/about-us/food-and-feed-codes-of-practice

Foods Standards Agency (2015). *Consistency in regulatory work.* http://food.gov.uk/research/research-projects/consistency-in-regulatory-work

Foods Standards Agency (2018). Regulating our Future. https://www.food.gov.uk/about-us/regulating-our-future

Food Standards Agency (2020). Assessing the contribution made by the food chain to the burden of UK-acquired norovirus infection. https://www.food.gov.uk/sites/default/files/media/document/assessing-the-contribution-made-by-the-food-chain-to-the-burden-of-uk-acquired-norovirus-infection.pdf

Garrie, D. and Keeler, A. (1993). Incomplete Enforcement with Endogenous Regulatory Choice. Discussion Paper 873, Institute for Economic Research, Queen's University, Canada.

Glachant, J.M., Khalfallah, H., Perez, Y., Rious, V. and Saguarn, M. (2013). Implementing Incentive Regulation and Regulatory Alignment with Resource Bounded Regulators. European University Institute Working Paper RSCAS 2013/31, Florence School of Regulation.

Glicksman, R.L. and EarnHart, D.H. (2007). Depiction of the Regulator-Regulated Entity Relationship in the Chemical Industry: Deterrence-Based vs. Cooperative Enforcement. *William and Mary Environmental Law and Policy Review*, Vol. 31, No. 3, 603-660.

Government of Canada (2018). Policy on Limiting Regulatory Burden on Business. https://www.canada.ca/en/treasury-board-secretariat/services/federal-regulatory-management/guidelines-tools/policy-limiting-regulatory-burden-business.html

Government of the Netherlands (2017). Better Regulation: Towards a Responsible Reduction in the Regulatory Burden Between 2012 – 2017. Ministry of Economic Affairs. https://www.government.nl/documents/reports/2017/08/22/better-regulation-towards-a-responsible-reduction-in-the-regulatory-burden-2012-2017

Gunningham, N. (2011). Enforcing Environmental Regulation. *Journal of Environmental Regulation*, Vol. 23, No. 2, 169-201.

Gunningham, N. (2017). Compliance, Enforcement, and Regulatory Excellence. RegNet Research Paper No. 124, School of Regulation and Global Governance (RegNet).

Hawkins, K. (1984). *Environment and Enforcement: Regulation and the Social Definition of Pollution*. Clarendon: Oxford.

Hawkins, K. (2002). *Law as Last Resort: Prosecution Decision-Making in a Regulatory Agency*. Oxford University Press: Oxford.

Hawkins, K. and Hutter, B.M. (1993). The Response of Business to Social Regulation in England and Wales: An Enforcement Perspective. *Law & Policy*, Vol. 15, No. 3, 199-217.

Ho, D.E., Ashwood, Z.C. and Handan-Nader, C. (2019). New Evidence on Information Disclosure Through Restaurant Hygiene Grading. *American Economic Journal: Economic Policy*, Vol. 11, No. 4, 404-428.

Hill, M. and Hupe, P. (2002). *Implementing Public Policy: Governance in Theory and in Practice*. Sage Publications: London.

Hutter, B.M. (1988). *The Reasonable Arm of the Law? Law Enforcement Procedures of Environmental Health Officers*. Clarendon: Oxford.

Hutter, B.M. (1989). Variations in Regulatory Enforcement Styles. *Law & Policy*, Vol. 11, No. 2, pp. 153-174.

Hutter, B.M. (1997). *Compliance: Regulation and Environment*. Clarendon: Oxford.

Jin, G.Z. and Leslie, P. (2003). The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards. *The Quarterly Journal of Economics*, Vol. 118, Issue 2, 409-451.

Jin, G.Z. and Leslie, P. (2009). Reputational Incentives for Restaurant Hygiene. *American Economic Journal: Microeconomics*, Vol. 1, No. 1, 237-267.

Kagan, R.A., Gunningham, N. and Thornton, D. (2003). Explaining Corporate Environmental Performance: How Does Regulation Matter? *Law & Society Review*, Vol. 37, No. 1, 51-90.

Khanna, V.S. (2021). Compliance as Costs and Benefits. In: *Cambridge Handbook of Compliance* (Eds. Van Rooij, B. and Sakol, D.D.), Part 1, Chapter 2, pp. 13-26. Cambridge University Press: Cambridge.

Kuperan, K. and Sutinan, J.G. (1998). Blue water crime: Deterrence, legitimacy, and compliance in fisheries. *Law & Society Review*, Vol. 32, No. 2., 309-338.

Lipsky, M. (1980). *Street level Bureaucracy: Dilemmas of the Individual in Public Service*. Russel Sage Foundation: New York.

Luukkanen, J., Nevas, M., Fredriksson-Ahomaa, M. and Lundén, J. (2018). Developing Official Control in Slaughterhouses Through Internal Audits. *Food Control*, Vol. 90, pp. 344-351.

Mackay, M., Yamazaki, S., Jennings, S., Sibly, H., Putten, I.E. and Emery, T.J. (2020). The influence of nudges on compliance behavior in recreational fisheries: A laboratory experiment. *ICES Journal of Marine Science*. Vol. 77, No. 6, 2319-2332.

Malesky, E. and Taussig, M. (2017). The Danger of Not Listening to Firms: Government Responsiveness and the Goal of Regulatory Compliance. *Academy of Management Journal*, Vol. 60, No. 5, 1741-1770.

Markel, D.L. (2000). The Role of Deterrence-Based Enforcement in a 'Reinvented' State/Federal Relationship: The Divide Between Theory and Reality. *Harvard Environmental Law Review*, Vol. 24, 1-114.

Markel, D.L. (2005). 'Slack' in the Administrative State and its Implications for Governance: The Issue of Accountability. *Oregon Law Review*, Vol. 84, No. 1, 1-65.

May, P.J. (2005). Compliance Motivations: Perspectives of Farmers, Homebuilders, and Marine Facilities. *Law & Policy*, Vol. 27, No. 2, 317-347.

May, P.J. and Winter, S.C. (2011). Regulatory Enforcement Styles and Compliance. In Parker, C. and Nielsen V.L. (Eds.), *Explaining Regulatory Compliance: Business Responses to Regulation* (pp. 232-244). Edward Elgar: Cheltenham.

May, P.J. and Wood, R.S. (2003). At the Regulatory Front Line: Inspector's Enforcement Styles and Regulatory Compliance. *Journal of Public Administration Research and Theory*, Vol. 13, No. 2, 117-139.

Meyers, M.K. and Vorsanger, S. (2003). Street-Level Bureaucrats and the Implementation of Public Policy. In (Eds. Peters, G. and Pierre, J.) *Handbook of Public Administration* (pp.153-164). Sage Publications: London.

Mintz, J.A. (1995). *Enforcement at the EPA: High Stakes and Hard Choices*. University of Texas Press: Austin.

Nagin, D.S. (2013). *Deterrence in the Twenty-First Century. Crime and Justice*, Vol. 42, No. 1, 199-263.

Nevas, M., Kalenius, S. and Lundén, J. (2013). Significance of Official Food Control in Food Safety: Food Business Operators' Perceptions. *Food Control*, Vol. 31, No. 1, 59-64.

Oded, S. (2013). *Corporate Compliance: New Approaches to Regulatory Enforcement*. Edward Elgar: Cheltenham.

OECD (2005). *Assuring Environmental Compliance: A Toolkit for Building Better Environmental Inspectorates in Eastern Europe, Caucasus and Central Asia*. OECD Publishing: Paris.

OECD (2013). *Cooperative compliance: A framework. From enhanced relationship to cooperative compliance*. Paris, OECD Publishing.

Patel, P.C. and Rietveld, C.A. (2020). The Impact of the Public Disclosure of Curved Inspection Scores Using Emojis on Hygiene Violators in Food Establishments. *Cornell Hospitality Quarterly*, 1-13. https://journals.sagepub.com/doi/10.1177/1938965520935398

Pautz, M.C. (2009). Trust between Regulators and the Regulated: A Case Study of Environmental Inspectors in Virginia. *Politics & Policy*, Vol. 37, No. 5, 1047-1072.

Pautz, M.C. (2010). Front-line Regulators and their Approach to Environmental Regulation in Southwest Ohio. *Review of Policy Research*, Vol. 27, No. 6, 761-780.

Pearce, F. and Tombs, S. (1997). Law and Class: Contextualizing the Regulation of Corporate Crime. *Social & Legal Studies*, Vol. 6, No. 1, 79-107.

Peters, D. H., Adam, T., Alonge, O., Akun, A. and Tran. N. (2013). Implementation Research: What It Is and How To Do It. *British Medical Journal*, BMJ 347: f6753. https://pubmed.ncbi.nlm.nih.gov/24259324/

Rechtschaffen, C. (1998). Deterrence vs. Cooperation and the Evolving Theory of Environmental Enforcement. *Southern California Law Review*, Vol. 71, No. 6, 1181-1272.

Rechtschaffen, C. (2004). Promoting Pragmatic Risk Regulation: Is Enforcement Discretion the Answer? *Kansas Law Review*, Vol. 53, 1327-1375.

Rechtschaffen, C. and Markell, D.L. (2003). *Reinventing Environmental Enforcement and the State/Federal Relationship*. Environmental Law Institute: Washington.

Regulating Futures Review (2017). Cabinet Office, UK Government. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/582283/Regulatory_Futures_Review.pdf

Russel, C.S., Harrington, W. and Vaughan, W.J. (1986). *Enforcing Pollution Control Laws*. John Hopkins University Press: Baltimore.

Sabatier, P.A. (1986). Top-Down and Bottom-Up Approaches to Implementation Research. *Journal of Public Policy*, Vol. 6, No. 1, 21-48.

Serapiglia, T., Kennedy, E., Thompson, S. and de Burger, R. (2007). Association of Food Premises Inspection and Disclosure Program with Retail-Acquired Foodborne Illness and Operator Noncompliance in Toronto. *Journal of Environmental Heath*, Vol. 70, No. 1, 54-59.

Scholz, J.T. (1984). Cooperation, Deterrence, and the Ecology of Regulatory Enforcement. *Law & Society Review*, Vol. 18, No. 2, 179-224.

Scholz, J.T. (1997). Enforcement Policy and Corporate Misconduct: The Changing Perspective of Deterrence Theory. *Law and Contemporary Problems*, Vol. 60, No. 3, 253-268.

Scholz, J.T. and Gray, W.B. (1997). Can Government Facilitate Cooperation? An Informational Model of OSHA Enforcement. *American Journal of Political Science*, Vol. 41, No. 3, 693-717.

Simon, P., Leslie, P., Jin, G.Z., Reporter, R., Aquirre, A. and fielding, J.E. (2003). Impact of Restaurant Hygiene Grade Cards on Foodborne-Decease Hospitalizations in Los Angeles County. *Journal of Environmental Heath*, Vol. 67, No. 7, 32-56.

Stigler, G.J. (1970). The Optimum Enforcement of Laws. *Journal of Political Economy*, Vol. 78, No. 3, 526-536.

Stoughton, M., Herb, J., Sullivan, J. and Crow, M. (2001). Toward Integrated Approaches to Compliance Assurance. *Environmental Law Reporter*, Vol. 31, 11266-11283.

Traill, W.B. and Koenig, A. (2010). Economic Assessment of Food Safety Standards: Costs and Benefits of Alternative Approaches. *Food Control*, Vol. 21, No. 12, 1611-1619.

Weske, U., Boselie, P., van Rensen, E.L.J. and Schneider, M.M.E. (2018). Using Regulatory Enforcement Theory to Explain Compliance with Quality and Patient Safety Regulations: The Case of Internal Audits. *BMC Health Services Research*, Vol. 18, No. 62, 1-6.

Wong, M.R., McKelvey, W., Schiff, C., Jacobson, J.B. and Kass, D. (2015). Impact of a Letter-Grade Program on Restaurant Sanitary Conditions and Diner Behavior in New York City. *American Journal of Public Health*, Vol. 105, No. 3, 81-87.

Yapp, C. and Fairman, R. (2005). Assessing Compliance with Food Safety Legislation in Small Businesses. *British Food Journal*, Vol. 107, No. 3, 150-161.

Yu, S. and Costanigro, M. (2019). The Effect of Public Online Disclosure of Restaurant Inspection Scores in the Lives Program: A Difference-in-Difference and Geographic Regression Discontinuity Approach. Annual Meeting of the Agricultural and Applied Economics Association, July 21-23, Atlanta, Georgia. file:///C:/Users/d74wi/Downloads/Abstracts_19_05_15_22_18_48_70__129_19_63_105_0.pdf

For the Editor and referees: a note on the appendices

The appendices have been written to provide full technical explanations and interpretations of the econometric models for (i) editors and referees and (ii) interested readers. To save journal pages, the appendices can be provided as on-line supplements. The tables are from Stata log files and are thus tamper-proof photographic evidence of the reported estimates. Moreover, including both the variable names as they appear in our dataset together with the syntax used to estimate the model enables easy replication and/or modification of our model by other researchers. Once refereeing is complete, the tables will be reset in a conventional publication format. Finally, for purposes of refereeing, in the three appendices the four local authorities are not anonymized.

Appendix A. Source of Figure 2a: Econometric model estimating food hygiene and safety compliance

The estimated model reported in Table A1 (below) explains variation in "Overall Risk" scores (the dependent variable) by variations in the variables of interest (the explanatory or independent variables): Inspection-Authority effects in each year (our main variables of interest); and the Business Category (although only one category – "Restaurants and Caterers" – varies over time and can thus be included in the model).

In Table A1 below, the Inspection-Authority effects in each year are captured by the estimated effects of the following authority-year dummy variables:

- DV_Cannock_09_10 is a dummy (binary) variable defined as one for all observations on inspections undertaken by the Cannock local authority (LA4 in Figure 1 and thereafter) in the year 2009-10 and as zero for all other observations. Similar dummy variables are defined for all observations on inspections undertaken by the Cannock local authority in each subsequent year in the dataset: i.e., DV_Cannock_10_11 … DV_Cannock_17_18.
- Similar sets of dummy variables for each of the remaining three local authorities: Newcastle or LA3 (DV_Newcastle_09_10 … DV_Newcastle_17_18); Stafford or LA2 (DV_Stafford_09_10 … DV_Stafford_17_18); and Stoke or LA1 (DV_Stoke_09_10 … DV_Stoke_17_18).

The model reported in Table A.1 also includes a business category dummy variable of all food handling establishments in the "Restaurants and Caterers" category (new_6 = 1; = 0 for all inspections of establishments in other business categories).

The model includes a “fixed effect” for each establishment, which controls for all constant (or slowly moving) but unobserved characteristics of the establishment. In our model, although the local authority effects are time invariant and thus potentially absorbed by the fixed effects, we interact them with year effects to create time-varying local authority effects that are consistent with fixed effects estimation. Hence, regarding the specification of our fixed effects model, separate sets of local authority and year dummies cannot be included, because the local authority dummies are fully absorbed by the fixed effects, and the year dummies are absorbed by the authority-year dummies, which are our variables of interest. Local authority dummies – with the omission of one as the base category – can be included in a random effects model; however, the Hausman test unambiguously rejects random effects estimation in favour of fixed effects estimation (p=0.0000).

The results from estimating the model reported in Table A.1 (below) tell us that the variables of interest explain 6.6 percent of the variation in establishments’ Overall Risk scores over time (R-square within). However, the model also accounts for variation in establishments’ Overall Risk scores by taking into account (controlling for):

- the fixed effects (sigma_u), which account for 85 percent of the variation in Overall Risk ratings that is not explained by the Authority-Year effects and the Business Category effect; and
- purely random events affecting scores (sigma_e), which account for 15 percent of the variation in Overall Risk ratings that is not explained by the variables of interest.

Of the variables of interest, holding all other influences constant, “Restaurants and caterers” (new_6) on average have worse (i.e., higher) Overall Risk ratings than all other categories by 27.71 points (almost 60 percent of the mean of 46.21).

Finally, we interpret the estimated Authority-Year effects on the Overall Risk ratings. These effects are estimated by Authority for each Authority-Year relative to its own performance in the base year 2008-09 (+’ve indicates deterioration; -’ve indicates improvement). These Authority-Year effects are interpreted as follows.

- Cannock effects:
    - All $9 \neq 0$ (all statistically significant at the 1 percent level)
    - Substantial improvements of the Overall Risk score of between -8.92 and -19.03
- Newcastle effects:
    - 7 from $9 \neq 0$ (statistically significant at the 1 percent level)
    - Small statistically significant improvements of between -3.75 and -7.10
- Stafford effects:
    - 7 from $9 \neq 0$ (statistically significant at the 1 or 5 percent level)
    - Small statistically significant improvements of between -1.35 and -3.39
- Stoke effects:
    - 8 from $9 \neq 0$ (statistically significant at the 1 percent level)
    - Small statistically significant improvements of between -1.90 and -6.13

These are the Authority-Year effects depicted graphically in Figure 1.

Table A1. Food Hygiene and Safety Compliance by Local Authority: Changes in average Overall Risk score from 2008-09 to 2017-18 (each year's estimate is relative to the base year of 2008-09) (fixed effects estimation with cluster-robust standard errors)

```
. xtreg Overall_Risk  DV_Cannock_09_10-DV_Cannock_17_18  DV_Newcastle_09_10-DV_Newcastle_17_18   DV_Stafford_09_10-DV_Stafford_17_18 DV_Stoke_09_10-DV_Stoke_17_18 new_6 , fe vce(cluster Business_ID_num)

Fixed-effects (within) regression               Number of obs     =     26,285
Group variable: Business_I~m                    Number of groups  =     11,294

R-sq:                                           Obs per group:
     within  = 0.0656                                         min =          1
     between = 0.0723                                         avg =        2.3
     overall = 0.0664                                         max =          6

                                                F(36,11293)       =          .
corr(u_i, Xb)  = -0.3929                        Prob > F          =          .

                          (Std. Err. adjusted for 11,294 clusters in Business_ID_num)
```

Overall_Risk	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
DV_Cannock_09_10	-8.918968	1.31874	-6.76	0.000	-11.50393	-6.334009
DV_Cannock_10_11	-7.757392	1.234305	-6.28	0.000	-10.17685	-5.337939
DV_Cannock_11_12	-10.12385	1.281709	-7.90	0.000	-12.63623	-7.611481
DV_Cannock_12_13	-11.43068	1.290281	-8.86	0.000	-13.95985	-8.901505
DV_Cannock_13_14	-12.38715	1.470415	-8.42	0.000	-15.26942	-9.504885
DV_Cannock_14_15	-19.03334	1.550218	-12.28	0.000	-22.07204	-15.99465
DV_Cannock_15_16	-15.98279	1.275873	-12.53	0.000	-18.48372	-13.48186
DV_Cannock_16_17	-17.10126	1.497388	-11.42	0.000	-20.0364	-14.16612
DV_Cannock_17_18	-15.21957	1.338399	-11.37	0.000	-17.84306	-12.59607
DV_Newcastle_09_10	.6121218	1.203276	0.51	0.611	-1.746509	2.970752
DV_Newcastle_10_11	-1.478202	1.047135	-1.41	0.158	-3.530769	.5743644
DV_Newcastle_11_12	-3.751696	.8319671	-4.51	0.000	-5.382496	-2.120896
DV_Newcastle_12_13	-5.194551	1.002734	-5.18	0.000	-7.160084	-3.229018
DV_Newcastle_13_14	-5.18848	1.02948	-5.04	0.000	-7.20644	-3.170521
DV_Newcastle_14_15	-5.98618	1.030483	-5.81	0.000	-8.006106	-3.966255
DV_Newcastle_15_16	-5.668974	.9670045	-5.86	0.000	-7.564471	-3.773477
DV_Newcastle_16_17	-6.156333	1.012542	-6.08	0.000	-8.141091	-4.171575
DV_Newcastle_17_18	-7.096227	1.081413	-6.56	0.000	-9.215984	-4.97647
DV_Stafford_09_10	.0792558	.854806	0.09	0.926	-1.596313	1.754824
DV_Stafford_10_11	.8842524	.7111884	1.24	0.214	-.5098007	2.278305
DV_Stafford_11_12	-1.35663	.6420245	-2.11	0.035	-2.61511	-.0981503
DV_Stafford_12_13	-1.477971	.6220996	-2.38	0.018	-2.697394	-.2585473
DV_Stafford_13_14	-2.916624	.6925139	-4.21	0.000	-4.274072	-1.559176
DV_Stafford_14_15	-3.393236	.6408109	-5.30	0.000	-4.649337	-2.137135
DV_Stafford_15_16	-2.865595	.744376	-3.85	0.000	-4.324702	-1.406489
DV_Stafford_16_17	-3.911423	.7637302	-5.12	0.000	-5.408467	-2.414379
DV_Stafford_17_18	-3.486321	.7199078	-4.84	0.000	-4.897466	-2.075177
DV_Stoke_09_10	-.0921684	.5073731	-0.18	0.856	-1.086708	.9023713
DV_Stoke_10_11	-1.899822	.5236085	-3.63	0.000	-2.926186	-.8734579
DV_Stoke_11_12	-3.885271	.5729264	-6.78	0.000	-5.008307	-2.762235
DV_Stoke_12_13	-5.527211	.5820055	-9.50	0.000	-6.668043	-4.386379
DV_Stoke_13_14	-5.725727	.56933	-10.06	0.000	-6.841713	-4.609741
DV_Stoke_14_15	-4.91942	.6365144	-7.73	0.000	-6.167098	-3.671741
DV_Stoke_15_16	-5.984572	.6411776	-9.33	0.000	-7.241392	-4.727753
DV_Stoke_16_17	-6.13229	.6452376	-9.50	0.000	-7.397068	-4.867512
DV_Stoke_17_18	-3.83727	.4241652	-9.05	0.000	-4.668707	-3.005832
new_6	27.71053	.6974168	39.73	0.000	26.34347	29.07759
_cons	30.70762	.6095552	50.38	0.000	29.51278	31.90245
sigma_u	20.420963					
sigma_e	8.7443782					
rho	.84505113	(fraction of variance due to u_i)				

Appendix B.

Table B1. Descriptive statistics for the sample used to estimate the model reported in Table A.1

Overall_Risk is the dependent variable, which is continuous over the range 0 to 147 with mean 46.21 and standard deviation 19.87.

DV_Cannock_10 is a truncation of DV_Cannock_09_10 (and similarly for DV_Newcastle_09_10 … DV_Stoke_17_18); for each of these dummy variables, the mean is the corresponding proportion of observations (e.g. inspections taking place in Cannock in the year 2009-10 account for 1.35% of the total number of observations used to estimate the model reported in Table A.1).

Estimation sample xtreg

Number of obs	=	26,285	
Number of clusters	=	11,294	
Obs per cluster: min	=	1	
avg	=	2.3	
max	=	6	

Variable	Mean	Std. Dev.	Min	Max
Overall_Risk	46.20677	19.87366	0	147
DV_Cannoc~10	.0135058	.1154292	0	1
DV_Cannoc~11	.0100818	.0999026	0	1
DV_Cannoc~12	.0127069	.1120084	0	1
DV_Cannoc~13	.016207	.1262731	0	1
DV_Cannoc~14	.009435	.0966766	0	1
DV_Cannoc~15	.0070002	.0833754	0	1
DV_Cannoc~16	.0190603	.1367396	0	1
DV_Cannoc~17	.0118318	.1081309	0	1
DV_Cannoc~18	.0141526	.1181219	0	1
DV_Newcas~10	.0065817	.0808618	0	1
DV_Newcas~11	.0090165	.0945283	0	1
DV_Newcas~12	.0485828	.2149984	0	1
DV_Newcas~13	.0213049	.1444016	0	1
DV_Newcas~14	.0210006	.1433887	0	1
DV_Newcas~15	.0205821	.1419832	0	1
DV_Newcas~16	.0229028	.1495965	0	1
DV_Newcas~17	.0193266	.1376729	0	1
DV_Newcas~18	.0184516	.1345801	0	1
DV_Staffo~10	.008522	.0919221	0	1
DV_Staffo~11	.0140765	.1178085	0	1
DV_Staffo~12	.0208103	.1427517	0	1
DV_Staffo~13	.0433327	.2036088	0	1
DV_Staffo~14	.0254137	.1573811	0	1
DV_Staffo~15	.025604	.1579536	0	1
DV_Staffo~16	.0262888	.1599957	0	1
DV_Staffo~17	.0252235	.1568063	0	1
DV_Staffo~18	.0311585	.1737491	0	1
DV_Stoke_~10	.0369412	.1886211	0	1
DV_Stoke_~11	.0325661	.1775014	0	1
DV_Stoke_~12	.0319193	.1757887	0	1
DV_Stoke_~13	.0253376	.1571515	0	1
DV_Stoke_~14	.031615	.1749761	0	1
DV_Stoke_~15	.0262507	.159883	0	1
DV_Stoke_~16	.0311965	.1738517	0	1
DV_Stoke_~17	.0322237	.1765971	0	1
DV_Stoke_~18	.1852768	.3885293	0	1
new_6	.727373	.445319	0	1

Std. Dev. not adjusted for clustering

Appendix C: The *Resource* variable

Table C1: *Resource* by Authority and Year (= FTE / Number of inspections) (rounded)

	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18
LA4	0.034	0.006	0.007	0.009	0.007	0.012	0.016	0.006	0.010	0.008
LA3	0.022	0.023	0.017	0.003	0.007	0.007	0.007	0.007	0.008	0.008
LA2	0.014	0.013	0.010	0.006	0.003	0.005	0.005	0.005	0.004	0.003
LA1	0.011	0.008	0.009	0.008	0.009	0.006	0.007	0.006	0.005	0.001