

Evaluating workplace inspections

Lars Lindblom BA, Royal Institute of Technology, Stockholm, and Sven Ove Hansson PhD, Royal Institute of Technology, Stockholm

Abstract

This paper introduces a theoretical framework for the evaluation of workplace inspections with respect to their effects on working conditions. The choice of a concept of efficiency is discussed, and its relation to criteria for a good working environment is clarified. It concludes that in order to obtain reliable information on the effects of different inspection methods, it is necessary to perform controlled comparative studies in which different methods are used in different workplaces. Given the ease with which such studies can be performed, it is surprising how few have been made. The studies that are available provide sufficient evidence that inspections can increase compliance with regulations and that they can also increase workplace safety, but not much can be concluded about the relative efficiency of different inspection methods.

Key words

Compliance, efficiency, enforcement, evaluation, inspection, priority-setting, quality of work, work environment, workplace health and safety, workplace inspection

Introduction

Although workplace inspections are by no means the only available method of reducing risks to human health in workplaces, their central role in current practice is indisputable. Therefore, the efficiency of workplace inspections is an essential issue in workplace health and safety. However, surprisingly little is known about whether (and then to what degree) inspections and other enforcement activities have the desired effects on working conditions:¹

... little research exists on the actual regulatory experience: that is, how enforcement encourages compliance and how compliance behaviour influences enforcement allocation.

Only fragmentary knowledge is available about the effects of enforcement on compliance behaviour and working conditions. The field is characterised by both a lack of reliable empirical data and a lack of theoretical underpinnings for the interpretation of such data.

The purpose of the present contribution is to provide a theoretical framework for the evaluation of workplace inspections, based both on international literature and on experiences from the evaluation of Swedish inspection agencies.²⁻⁴ The second section is devoted to the choice of evaluation criteria and efficiency measures. This is followed by a discussion of what evaluation methodology should be used, and a section in which results from previous research are summarised and assessed. In the concluding section the need for more controlled comparative studies of inspection methods is emphasised, and it is pointed out that workplace inspection agencies have excellent opportunities to improve their efficiency by systematic testing of different inspection methods.

Evaluation criteria

An evaluation of a social measure can be defined as a study that attempts to find out to what degree certain goals (criteria, standards) for that measure have been satisfied. Since evaluations

aim at determining goal satisfaction, they have a close relation to the notion of efficiency. Efficiency means satisfaction of goals (criteria, standards). To determine the degree to which goals have been reached means to determine the degree of efficiency in the satisfaction of these goals.

Efficiency must always be defined in relation to a set of goals (explicit or implicit). In some cases, reference is made to only a single goal, and efficiency (in this case more commonly called 'effectiveness') means that this goal is satisfied to as high a degree as possible. In most cases, however, efficiency refers to two or more goals. In such instances, efficiency has been achieved with respect to the given goals if, and only if, none of these goals could have been achieved to a higher degree without some of the others being achieved to a lower degree.⁵

Discussions of efficiency are often confused by a lack of precision in determining the dimensions of the goal. In most contexts, the two most important notions of efficiency are effectiveness and cost-efficiency. A measure is *effective* to the extent that it solves the problem that it is intended to solve. Effectiveness has nothing to do with costs. An effective measure may very well be indefensible because it consumes too many resources, resources that could have been better used. A measure is *cost-efficient* to the extent that it is efficient with respect to two goals: improvement and cost minimisation. More precisely, it is cost-efficient if, and only if, no alternative exists that achieves more effect at the same or lower price, or less expenditure for the same or better effect. A cost-efficient alternative may very well be indefensible, either because it produces too little effect (albeit at low cost) or because it costs too much.

In the evaluation of inspections, there are two types of goals that have to be taken into account: first, the goals that refer to conditions in the workplace and to quality of work life in general; second, the goals that take costs into account. The former type of goals will be discussed in the next section, and their combination with economic goals in the section on criteria of efficiency.

The work quality dimensions

An obvious success criterion for workplace inspections is the degree to which they lead to increased compliance with regulations. However, compliance as such is not in general the intended final outcome. There is an underlying assumption that rule compliance will lead to more concrete effects such as lowered injury rates and other improvements in working conditions. Although this assumption is in most cases reasonable, the step from rule compliance to improved conditions is not in all cases automatic. It may be fairly automatic for a regulation about machine safety, but far less so in the case of a regulation about the internal health and safety organisation of the regulated businesses. As can be seen, for instance, from Hutter's⁶ in-depth study of occupational safety and health in the British railway system, the relationship between rule-setting, compliance with rules and actual workplace conditions is complex and may vary in different types of businesses and workplaces.

The model shown in Figure 1 is proposed as a conceptual framework for the relationship between rule compliance and improvements in working conditions. The major intended mechanism by which enforcement reduces risks is through increased compliance (arrows 1 and 2). In addition, enforcement can also have direct effects, not mediated by compliance with the written regulations (arrow 3). For instance, inspections may lead to improved safety routines or heightened health and safety consciousness in a firm. Boyne *et al.*⁷ have identified important factors that can impede the effectiveness of inspection. For a deeper empirical understanding of the effects of enforcement, each of the three links in Figure 1 should be studied as much as possible in isolation. This means

that, if possible, the effects of inspections should be studied both in terms of rule compliance and in terms of indicators of the quality of actual working conditions.

In order to determine whether or not workplace conditions have improved, we need standards or criteria for the satisfactoriness of working conditions, ie essentially a normative theory for what work should be like. There are many traditions of thought in this area, with widely divergent ideas. In order to give an impression of this diversity, five different traditions of theories of work will be surveyed: philosophical theories, theories from occupational safety and health, management theory, sociological theories, and business ethics.

The most influential philosophical tradition with respect to work is the Aristotelian one, according to which, for work to be characterised as good, it must be conducive to individual flourishing or self-realisation. This view has re-emerged in many variants. Marx and his followers have often referred to self-realising work as the antithesis of the alienated work that dominates in modern societies.^{8,9} Writing in another tradition, influenced by writers such as Tolstoy and Thoreau, Gandhi argued that work should be seen as an expression of our being.¹⁰ In Pope John Paul II's famous encyclical *Laborem Exercens* ('On human work'),¹¹ the value of work is said to be determined by its conduciveness to human development. The value of work is 'judged above all by *the measure of the dignity* of the subject of work, that is to say the person, *the individual who carries it out.*' [authors' emphasis] In summary, the following criterion is indicated by many influential philosophers of work:

- Good work leads to self-realisation.

Occupational safety and health is a multi-disciplinary endeavour, including a diversity of disciplines 'working toward a common goal, ie to maintain a safe and healthful workplace'.¹² A recent textbook starts its introduction by saying that 'the field of occupational safety has been broadened to encompass both safety and health'.¹³ There are two partially overlapping ways to look at health and safety:

One view is that safety is concerned with injury-causing situations, whereas health is concerned with disease-causing conditions. Another view is that safety is concerned with hazards to humans that result from sudden severe conditions; health deals with adverse reactions to prolonged exposure to dangerous, but less intense, hazards.¹³

For this discipline the overriding quality criterion for work is to avoid disease and accidents. This implies the following criterion:

- Good work is healthy and safe.

The founding father of management theory, Peter Drucker, warns against 'the fallacy of the dominant dimension'.¹⁴ This fallacy consists in conflating the different dimensions of work.

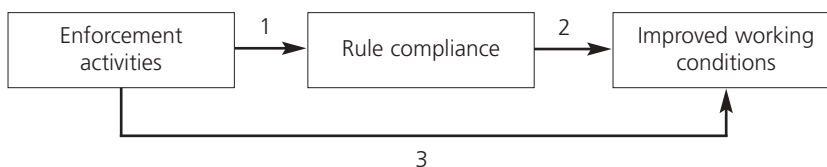


Figure 1
A framework for studies of efficiency in enforcement

There are six dimensions according to Drucker. The physiological dimension concerns the need for variation in work. The psychological dimension relates to work as an extension of personality. Work can also be seen as a social bond, in that it is a means to satisfy our need for belonging to a group and having meaningful relationships with other people. But work is also the means to earn a living, which leads on to the economic dimension. There are two aspects of this dimension: the individuals' earning of a living and the 'earning of a living' of the company. According to Drucker there is a necessary conflict between the two. The next two dimensions concern power. Organised work necessarily leads to the need for authority. It is an inherent fact of organisational life and an essential dimension of work that somebody must lead it. There is also a redistributive dimension of work. Since it is impossible to measure how much any one employee has contributed to an organisation's results, there is no objective standard for wages, and therefore authority once again comes into play. Drucker maintains that there is no overarching principle that can decide which is more important among these dimensions; it is a question of dynamics. It is surprising that Drucker nevertheless seems to commit the fallacy of the dominant dimension. After discussing these various aspects of work, he goes on to say:¹⁴

Nevertheless the manager has to manage now. He has to find solutions – or at least accommodations – which will enable him to make work productive and the worker achieving.

For Drucker there still seems to be a dominant dimension: productivity. As far as can be seen, in this respect he is representative of his discipline (see, for instance, Kinlaw¹⁵ and Hammer & Champy¹⁶). Management theory is multi-faceted and often includes several aspects of work. The characteristic feature that distinguishes it from the other disciplinary approaches to work under discussion here is the following:

- Good work is productive.

The sociology of work includes a research tradition that addresses directly the quality of working life (QWL). According to Walton,¹⁷ this concept consists of:

- adequate and fair compensation
- healthy and safe working conditions
- immediate opportunity to use and develop human capacities
- future opportunity growth and security
- social integration in the work organisation
- constitutionalism in the work organisation
- work and the total life space
- the social relevance of work life.

This is a very inclusive idea about work; Walton recognises that:¹⁷

The question arises whether there are inherent trade-off relationships among some of these qualities of working life which necessitate a decline in one quality in order to improve another.

However, if one takes QWL to be the most general genre of the social science of work, one can also use it to derive a substantial criterion for a theory of good work. Sirgy *et al.* state in a recent paper:¹⁸

Although there is no formal definition of quality of working life (QWL), industrial psychologists and management scholars agree in general that QWL is a construct that deals with the wellbeing of employees.

The sociology of work is multifaceted, but if one is to extract one characteristic criterion of good work from it, the following is a reasonable choice:

- Good work promotes wellbeing.

Finally, business ethics must be considered. The focal point here is usually how to help managers deal with the ethical dimension of their tasks; work is not a central issue.¹⁹ When work is discussed, the focus tends to be on rights in the workplace. One listing²⁰ identifies rights to a healthy and safe working environment, job security, privacy, compensation for injury, participation, equality, pension, union membership, freedom from harassment and a living wage. Often a right to meaningful work is suggested. On one interpretation,²¹ in order to be meaningful, work has to be freely entered into. Furthermore, it has to allow the worker to exercise autonomy and to develop his or her rational capacities, to provide a living wage, to support his or her moral development, and to be organised in a non-paternalistic way. Central discussions in business ethics revolve around issues such as respect for autonomy, fairness and rights. This can be summarised under the (admittedly vague) heading of ‘justice’:

- Good work is organised in a just manner.

Hence, five central values in the discussion on good work have been identified, namely that good work leads to self-realisation, is healthy and safe, is productive, promotes wellbeing and is organised in a just manner. This is a very wide set of criteria, and the problem of how to combine them, or weigh them against each other, has no obvious solution. However, although all these criteria are relevant in a discussion of quality of work in general, it is far from self-evident that they should all be taken into account in evaluations of workplace inspections. First, some of them, such as self-realisation, may be extremely difficult to operationalise for inspection purposes. Second, it can reasonably be argued that some aspects of good work should be furthered by means other than inspections carried out by government agencies. Two of the five criteria mentioned above, namely health and safety and wellbeing, seem to be more central to inspection agencies than the others. Of course, even if deliberations are limited to these criteria, there is still a problem of multi-dimensionality. Health and safety is internally multi-dimensional since it contains many different aspects, such as noise, chemical exposure, climate and so on. However, the health-related criteria are relatively easy to compare and combine since analysts have reasonably similar intuitions about the severity of illnesses.²²

In spite of the many aspects of health and safety, it is in most cases not difficult to determine whether or not a change in a workplace leads to an improvement in terms of health and safety. To determine whether it is an improvement in terms of the totality of the five aspects enumerated above is much more difficult. This is an area that has been neglected in ethical theory and it is high time that practitioners and philosophers start working together to develop a much-needed normative theory of work.

For practical purposes, inspectors and their agencies will, of course, have to base their decisions on the criteria of working conditions that have been laid down in laws and regulations. However, the authors consider it inadvisable to dissociate basic discussions on

work quality criteria from inspection agencies, since the experiences of inspections should be important inputs to such discussions.

Criteria of efficiency

There is a close relationship between, on the one hand, the criteria of efficiency used in evaluation studies and, on the other, principles of priority-setting. The latter can be seen as applications of criteria of efficiency, used prior to the implementation of a measure in order to ensure that resources are used efficiently. In day-to-day work in inspection agencies, priority-setting principles are more often used and discussed than evaluation criteria. The two forms of efficiency criteria will therefore be treated in parallel, presenting four alternative notions of efficiency in both forms.

The following is a simple principle that focuses on workplace conditions and does not take costs into account:

(1) **Worst things first**

Priority-setting: the highest priority is assigned to the worst working conditions.

Evaluation: efficiency is achieved when the quality of the worst working conditions is maximised.

From the viewpoint of justice, strong arguments can be made for (1). The employees who are exposed to the worst working conditions have a reasonable claim to receive help from the responsible authorities before help is given to others with less serious problems. (1) can also be seen as an analogue to the difference principle that is one of the most characteristic elements of John Rawls's²³ theory of justice. According to that principle, the degree of justice in a society is determined by the living conditions of those who are worst off in that society. According to (1), the success of attempts to improve working conditions in a society is determined by the state of the worst working conditions in that society.

Until recently, the Swedish labour inspectorate used as a result indicator that 55 per cent of the inspections should lead to written injunctions.²⁴ This indicator urges inspectors to visit the workplaces where they expect to find the worst working conditions, and can therefore be seen as an attempt to implement priority principle (1).

From an economic point of view, (1) is far from unproblematic. Arguably, when used for priority-setting, this criterion can lead to the misuse of resources, since 'it is crucial to tackle not only issues that are important, but problems that are amenable to solution'.²⁵ The following is an alternative criterion that shares with (1) the property of focusing entirely on working conditions, and thus does not take costs into account:

(2) **Maximum total improvement**

Priority-setting: the highest priority is assigned to the alternative that maximises total improvement.

Evaluation: efficiency is achieved when total improvement is maximised.

Approach (2) differs from (1) in paying no attention to the distribution of health and safety problems in workplaces. The following schematic example can be used to clarify the difference: an inspector can choose between two activities, A and B. Activity A is aimed at a small company in which 50 persons are exposed to conditions such that 10 per cent of them

contract a certain disease. Activity A reduces this frequency from 10 per cent to 0 per cent. Activity B is aimed at a large company in which 5,000 persons are exposed to conditions such that 1 per cent of them contract the same disease. Activity B reduces this frequency from 1 per cent to 0 per cent. According to priority-setting principle (1), activity A should be given higher priority since an individual risk of 10 per cent is worse than one of 1 per cent. According to priority-setting principle (2), activity B should be preferred since it leads to a much higher reduction of total morbidity. Activity B is expected to save 50 persons from the disease whereas activity A is only expected to save 5 persons.

In its recent strategy paper, the UK Health and Safety Commission (HSC) made it clear that, in their view, inspection and other agency activities should be 'targeted where they can have the most impact'.²⁶ This is in fact a clear endorsement of the idea expressed in principle (2).

Next, criteria that take costs into account must be considered. In the context of measures taken by public agencies, the notion of costs referred to in cost-efficiency is ambiguous and should be further clarified. It may refer either to costs in terms of agency resources or in terms of total social resources. These two delimitations of costs are both relevant, and so are the two notions of cost-efficiency that they give rise to.

The following principle seems to have some influence in public agencies:

(3) **Best use of agency resources**

Priority-setting: the highest priority is assigned to the alternative that maximises the total improvement per agency resource input.

Evaluation: efficiency is achieved when the total improvement per agency resource input is maximised.

This is a two-dimensional efficiency criterion (cost-efficiency), in which only those costs are taken into account that pertain to the agency's own resources. In the agency's own deliberations, in which the resources available to the agency have to be taken as a given, fixed amount, priority-setting principles (2) and (3) will yield the same policy recommendations, although (3) but not (2) mentions costs explicitly. Several studies indicate that prioritisations are sometimes made in approximate accordance with principles (2) and (3). One study of air pollution control in the USA showed that plants with good historic compliance rates received more attention than plants with bad compliance history. As the study authors point out:¹ 'This would indicate a tendency for regulators to exert pressure on firms that are more likely to respond without costly legal battles.'

A series of studies conducted by David Weil²⁷⁻²⁹ shows that OSHA inspectors focus their attention on large unionised establishments. Similarly, in Norway, plants with strong local unions are inspected more often by the Labour Inspectorate than plants with weak local unions.³⁰ One possible explanation of these results is that unions facilitate the implementation of the inspector's injunctions, so that fewer agency resources have to be spent in these plants to obtain compliance with the inspectorate's injunctions. Another contributing explanation is that union representatives report problems to the inspectorate and ask for inspectors to visit the workplace. A third possible explanation is that agencies prefer to work with businesses in which negotiated compliance is achievable, and therefore do not give priority to plants that are less willing to comply.³¹

A different approach is to take not only agency resources but total social resources into account.

(4) Best use of social resources

Priority-setting: the highest priority is assigned to the alternative that maximises the total improvement per total social resource input.

Evaluation: efficiency is achieved when the total improvement per total social resource input is maximised.

Contrary to (3), approach (4) cannot be reduced for practical purposes to a one-dimensional measure. It requires that the agency continuously optimise its activities according to estimates of total costs. Unfortunately, total social costs are often very difficult to estimate, due to lack of information and to the unpredictability of technological and social change. Often agencies depend for these estimates on information from businesses that have more to lose than to gain from providing authorities with truthful and unbiased information on the costs in question.

Other principles for priority-setting and evaluation are, of course, also possible. Although it is common to accuse regulators of not setting the right priorities, critics often leave their audience ignorant of exactly what is the criterion against which they measure actual performance. Studies of priority-setting that take these basic distinctions into account would be valuable contributions.

Various arguments can be made for and against these and other efficiency criteria and priority-setting principles for workplace inspections. In most practical cases, however, the choice between these approaches to efficiency has limited importance since the most basic information needed for their application is lacking: it is not known to what extent inspections (and other measures taken by government agencies to improve working conditions) actually lead to improvements in the working environment.

Therefore, any serious attempt to determine the efficiency of workplace inspections must begin with this, the most basic issue. Before dealing with more sophisticated issues such as weighing costs against workplace improvements, it is necessary to determine the extent to which inspections have the desired effects on workplaces.

Designing evaluation studies

Ideally, it would be desirable to know the total effects on working conditions of the inspection activities of a country's health and safety authority. It would be useful to know, for instance, if and to what extent workplaces are improved and work-related illness is reduced when more resources are spent on government activities in this area. There are two methods that could, at first glance, be expected to provide such information, namely time series and international comparisons. On closer inspection, it turns out that neither can provide reliable information on the total effects of agency activities. Therefore, such information is in practice impossible to obtain.

Time series are of very limited use since the development of working conditions in a country is influenced by many factors other than agency activities. Therefore, even if it is found, for instance, that total accident rates have increased (or decreased) after a decrease (or increase) in agency resources, no conclusion can be drawn from this because of the influence of other factors such as technological changes and the general economic development.

For similar reasons it is very difficult to draw conclusions from international comparisons. There are many factors in addition to agency activities that can lead to differences between countries in the quality of workplace conditions. Some of these factors correlate in complex ways with agency resources, so that the effects of the latter are in practice impossible to

isolate. In addition, it is very difficult to compare the extent and the character of workplace inspections in different countries.

Therefore, on the whole, evaluating total effects is not feasible in practice. What can be done with hope of success is to find out the effects of single measures. Even this, however, is far from simple. The usual evaluation method, in which the state of the working environment before and after the implementation of measures is compared, is problematic for much the same reasons as the use of time series in the evaluation of total agency activities. Suppose, for instance, that a new inspection method, or intensified use of traditional inspections, is tried out in a particular branch of industry, and that a substantial improvement in working conditions is found to have taken place. This may, of course, depend on the inspection activities, but it may also depend on other factors such as voluntary measures that industry would have taken even in the absence of inspections.

By far the best way to isolate the effects of inspections from those of other causal factors is to perform studies in which different workplaces are treated differently with respect to inspections. This means that organisations are inspected with different methods, in accordance with a study design that has been constructed to detect differences in working conditions that may be the result of these differences in inspection.

It is important to note that the reason why the efficiency of single measures but not that of total agency activities can be determined is exactly this: controlled comparisons are feasible in the former case but not in the latter. To obtain results that are fully reliable from a scientific point of view, it is not sufficient to shift focus from total effects to effects of single measures. It is also necessary to implement stringent study designs in which the effects of inspections are isolated as far as possible from the effects of other causal factors.

In order to achieve this, evaluation studies must be thoroughly integrated into the work of inspection agencies. The current standard procedure is to design evaluation studies only after all decisions on the actual inspections have been made. This procedure makes the task of evaluators next to impossible. In order to make well-informed methodological development possible, evaluation must be integrated in the general planning of inspection activities. This means that (new and old) methodologies, as well as different intensities of inspection, should be distributed among businesses in ways that are compatible with the design requirements of evaluation studies.

Unfortunately, there is currently a trend in the opposite direction from this. It has been proposed that government agencies cannot evaluate themselves, and that therefore all evaluations should be performed by external evaluators. The idea that execution and evaluation should be kept organisationally apart is, of course, sensible in principle, but it does not work in practice in cases – such as the evaluation of workplace inspections – in which meaningful evaluation is only possible if the requirements of evaluation design are integrated in the general planning of the activities to be evaluated.

It is instructive to compare the epistemological conditions for evaluating workplace inspections with those for evaluating medical treatment. Clinical trials, in which treatment of patient groups is organised in ways that facilitate the evaluation of treatment effects, is the only scientifically accepted methodology. It would be a big step backwards to replace them with outside evaluators who evaluated treatments that had not been organised in this way. Such external evaluators would miss most of the effects that can be discovered in well-organised clinical trials.

Few other organisations have better opportunities than inspection agencies to determine the effectiveness of their activities by systematic testing and evaluation of alternative methods – provided that inspection activities are planned in accordance with the requirements of evaluation design. If there are, say, 100 companies in a certain branch of industry to be inspected, then they can be randomly divided into two groups of 50 that are inspected by different methods. Possible evaluation methods include studies of injury records and later inspections with uniform methodologies.

It is important in such evaluation studies to distinguish between (complete) inspection methods and aspects or elements of such methods. The term *inspection method* denotes the (total) way in which inspection is conducted, including the way in which it is influenced by different reactions from the inspected business. Hence, an inspection method does not have to be standardised in a way that excludes flexibility. Many inspection methods are traditional, i.e. they have grown out of practice and have not been written down or otherwise made explicit in a systematic fashion. However, there are also inspection methods that have been consciously developed and described in considerable detail. Such methods have often been developed in response to perceived problems with entrenched, traditional methods.

A methodological *element* is defined as an aspect of the way in which inspections are conducted, such as whether or not they are pre-announced, the extent to which informal agreements are used rather than formal ticket-writing and so on. Evaluations can focus either on methods or on methodological elements. In an evaluation of (complete) inspection methods, different businesses are inspected with different methods, and the improvements achieved are compared. An obvious advantage of this type of evaluation is that the results are easy to interpret and implement: methods shown to be inefficient should be replaced by methods shown to be more efficient. A disadvantage is that since methods differ in many respects, it is not possible to determine from such a comparison which of the many differences between two evaluated methods are the causes of the observed differences. This can limit the applicability of the results obtained to new situations.

The other type of evaluation consists in modifying an inspection method with respect to a particular methodological element. An example of this would be to perform a certain type of inspection with and without pre-announcement. As was pointed out to the authors by an anonymous referee of this journal, some elements can only be included in certain types of methods. Hence, unannounced inspections can be used in methods focusing on checking the status of the physical environment, but not in methods that require extensive interviews and discussions with personnel. Therefore, this element can only be evaluated in the context of inspections that focus on physical rather than organisational matters.

From the viewpoint of practical usefulness, evaluation studies of both these types should have a high priority, not least as a support for the development of new methods of inspection. Indeed, the HSC has committed itself to developing an ‘evidence-based interventions strategy’, including not only inspection methodology but also other interventions that can increase compliance, such as insurance, guidance, training and competence-building.²⁶ However, very few controlled studies of inspections have been performed that can provide reliable evidence for this purpose. Possible explanations of the lack of such studies are the absence of an evaluation culture in public agencies, the low status of evaluation work, and insufficient understanding in government organisations of how activities should be organised in order to be evaluable.

Previous research

Surprisingly few evaluation studies of inspections have been performed that satisfy the quality criteria outlined in the previous section. In this section, a brief overview will be made of studies that are relevant for conclusions on the efficiency of inspections. The overview will be divided into two parts, according to the distinction made above between effects on rule compliance and effects on working conditions.

Effects of inspections on compliance

Several US studies indicate that OSHA (Occupational Safety and Health Administration) inspections result in improved compliance among regulated establishments. One of the best of these studies was a longitudinal study of the custom woodworking industry. Compliance with machine-guarding standards was shown to increase substantially after inspections. In view of the small fines, this increase in compliance could not be explained as the result of profit maximisation (as would have been anticipated according to the expected utility model of law abidance). Hence, non-profit-maximising mechanisms seem to contribute to employers' compliance behaviour.²⁹

According to another study of OSHA inspections, enforcement actions against firms with 100 to 500 employees have a greater impact than enforcement directed against larger or smaller firms. Furthermore, this study indicates – again contrary to the expected utility model of law abidance – that the size of penalties is inconsequential. Smaller penalties had the same effects as larger ones, but the legal procedures for larger penalties took considerably more time for the inspectors.³² As noted by Gunningham,³³ the relationship between profit maximisation and workplace safety varies from business to business, and will depend on such factors as profit margin and relationships to employees.

In a large body of material based on OSHA inspections in the years 1972–1983, the effects of inspections were measured in terms of the number of citations on subsequent inspections. This study showed that the initial inspection of an establishment drastically reduced subsequent violations. (Citations were reduced by about 50 per cent and exposures above OSHA limits by 42 per cent.) In contrast, subsequent inspections had little effect on compliance:^{34, 35}

The results suggest that, on the margin, substantial gains could occur if inspection resources were reallocated from the intensive margin to the extensive margin of OSHA's inspection strategy.

It is interesting to note in this context that the intuitive opinions of inspectors on the efficiency of prosecution are not uniform. Hence, in a study by Hawkins,³⁶ half of the interviewed factory inspectors believed that prosecution has an impact on the behaviour of employers, whereas the other half did not. Unfortunately, there are insufficient research data to adjudicate between the two opinions. Although evidence is available that indicates the positive effects of corporate sanctions, there also seem to be contexts in which sanctions adversely affect compliance with the law.³⁷

A study of the enforcement of US air pollution regulation on steel plants likewise showed that inspections led to greater compliance. Furthermore, steel plants owned by single-plant firms were more likely to comply than plants owned by multi-plant firms. A possible explanation for this result is that:¹

... if a single-plant firm's senior management was more likely to live in or near the community where their plant was located, then their greater compliance rate may have arisen from a desire to avoid any embarrassment that non-compliance might involve.

Wilthagen summarises a follow-up study of inspections by the Dutch Labour Inspectorate:³⁸

In 55 percent of the cases the employer had met the wishes or demands of the inspector after the initial visit. In 10 percent the firm had complied after two or more visits. In 22 percent the problems had only partially been solved. In about 10 percent of all cases the employer had not (yet) taken the measures required by the inspector. On further examination it turned out that the Labour Inspectorate was notably successful when minor measures had been demanded. Major, that is structural, safety and health problems in firms are less easily influenced and solved by labour inspectors. The intervention of labour inspectors hence runs the risk of appearing [to be] of a non-lasting and non-substantial nature.

In a study of the so-called 'macro' method (a method for intensive systems inspections developed in the Malmö branch of the Swedish labour inspectorate), compliance with rules of internal control was investigated with extensive structured interviews with the local parties in the workplaces. An index (the IK or 'internal control' index) was used to summarise the level of internal control in the workplaces. Similar workplaces on which the macro method was not applied were used as a control group. The results indicate that the macro method gives rise to improved internal control on inspected workplaces. This study also illustrates how innovative and complex methods of inspection can be evaluated by using control group studies.³⁹

Effects of inspections on health and safety

The studies referred to in the previous section all concern the effects of inspections on compliance with regulations. Even more interesting is their effects on the final outcomes that these regulations aim at, such as lowered injury rates and decreased atmospheric concentrations of hazardous substances. Several studies have shown effects of OSHA inspections on injury rates.^{40,41} One study showed that for plants with 200 or more employees, OSHA citation activities substantially reduced the number of days lost due to injury. Furthermore, this study gives indication that plant-specific programmes (jointly administered with unions) were more efficient in reducing injuries than OSHA activities, but the limited statistics make this last conclusion uncertain.⁴²

A study of the impact of OSHA inspections between 1979 and 1985 showed a significant effect: inspections imposing penalties induced on average a 22 per cent decline in injuries in the inspected plant in the following few years. There was also a decrease in work day losses, ie 15 to 18 fewer lost work days per 1 inspector work day.⁴³ In another paper, the same authors estimate that a 10 per cent increase in OSHA enforcement activity reduces injuries by 1 per cent in large, frequently inspected firms.⁴⁴

A more recent study⁴⁵ compared the effect of inspections with those of consultations on claims rates for worker's compensation. A strong association was found between enforcement visits and a decrease in claims for compensation. Inspected worksites experienced a 25 per cent decline in such claims. No significant effects were found for consultations.

A couple of studies of environmental inspections point in the same direction. In one study, Environmental Protection Agency inspections of pulp and paper companies with respect to water pollution turned out to have a strong effect on both pollution levels and on the rates of compliance with the permit levels.⁴⁶

Conclusions

To summarise these studies of enforcement efficiency, there is sufficient evidence to propose that inspections can increase compliance with regulations and that they can also increase workplace safety. Single studies indicate that the first inspection of a workplace has larger effects than subsequent inspections, that the size of fines is inconsequential, and that effects are larger in middle-sized and single-owner plants. However, most of these results refer to US inspection agencies, and not all of them may be universally applicable. From a European point of view it must be observed that both the internal workings of businesses and the relationships between businesses and government agencies differ significantly between the USA and Europe, and that there are also important differences between different European countries. There is a great need for studies of the efficiency of workplace inspections in different European countries. This has been demonstrated very clearly in the Swedish context; since the differences in labour relations between Sweden and the USA are so large, the applicability of results obtained in one of these countries to the other country is always uncertain.

Ongoing changes in working life pose new challenges to inspection agencies. This applies not least to the introduction of teleworking and distance working⁴⁷ and to new technologies for monitoring and control of the workforce that may lead to serious infringements of privacy.⁴⁸ Inspection methods to deal with new problems in workplaces need to be developed and evaluated. Furthermore, because of the current trend towards self-regulation,^{49, 50} it is important to investigate the efficiency of different ways for inspection agencies to promote self-regulation.

It should also be noted that most studies evaluate inspections *per se*, and do not investigate specific inspection methodologies. Studies that compare the efficiency of different methods of conducting inspections are needed.

Workplace inspections should be efficient. In order to make them efficient, solid evidence is needed on the actual effects on workplace conditions of inspections with different inspection methodologies. The lack of such evidence is surprising, given the facility with which it could have been obtained. Few other organisations have better opportunities than inspection agencies to improve their efficiency by systematic testing and evaluation of alternative methods. With current economic pressures on government agencies, it is high time to make use of this opportunity to improve the quality of inspection work.

References

1. Gray W B and Deily M E. Compliance and enforcement: air pollution regulation in the US steel industry. *Journal of Environmental Economics and Management* 1996; 31 (1): 96–111.
2. Hansson S O and Lindberg A-K. *Utvärdering av Arbetsmiljöverkets haverikommission* ('Evaluation of the Accident Investigation Committee of the Swedish Work Environment Authority'). Arbetsmiljöverket, 2004 (in press).
3. Lindblom L, Clausen J, Edvardsson K, Hayenhielm M, Hermansson H, Nihlén J, Palm E, Rudén C, Wikman P and Hansson S O. *How agencies inspect. A comparative study of inspection policies in eight Swedish government agencies*. Statens Kärnkraftsinspektion Report 2003: 36.
4. Eriksson O. Evaluation of interventions in the government sector in Sweden: a need for change. *Policy and Practice in Health and Safety* 2004; 2 (1): 75–89.
5. LeGrand J. *Equity and choice. An essay in economics and applied philosophy*. London: Harper Collins Academic, 1991.

6. Hutter B M. *Regulation and risk: occupational health and safety on the railroads*. Oxford: Oxford University Press, 2001.
7. Boyne G, Day P and Walker R. The evaluation of public service inspection: a theoretical framework. *Urban Studies* 2002; 39: 1197–1212.
8. Braverman H. *Labor and monopoly capital: the degradation of work in the twentieth century*. New York: Monthly Review Press, 1976.
9. Elster J. Self-realization in work and politics. *Social Philosophy and Policy* 1986; 3: 97–126.
10. Prabhu J. Gandhi's philosophy of work and its contemporary relevance. *Concepts and Transformations* 1987; 2: 73–88.
11. John Paul II (1981) *Laborem Exercens*: 7. www.vatican.va/holy_father/john_paul_ii/encyclicals/documents/hf_jp-ii_enc_14091981_laborem-exercens_en.html.
12. Murthy L. Introduction. In: LaDou J. *Occupational health and safety*. Itasca: National Safety Council, 1994: 15.
13. Goetsch D. *Occupational safety and health for technologists, engineers, and managers*. Upper Saddle River: Prentice-Hall, 1999: 1.
14. Drucker P F. *Management: tasks, responsibilities, practices*. New York: Harper Business, 1993.
15. Kinlaw D C. *The practice of empowerment – making the most of human competence*. Gower Publishing, 1995.
16. Hammer M and Champy J. *Reengineering the corporation – a manifesto for business revolution*. Nicholas Brealey Publishing, 1998.
17. Walton R E. Quality of working life: what is it? *Sloane Management Review* 1973; Fall: 11–21.
18. Sirgy M J, Efraty D, Siegel P and Lee D-J. A new measure of quality of work life (QWL) based on need satisfaction and spillover theories. *Social Indicators Research* 2001; 55 (3): 241–302.
19. Ciulla J B. Business ethics and work: questions for the twenty-first century. In: Fredrick R E (ed). *A companion to business ethics*. Oxford: Blackwell Publishers, 1999.
20. Duska R. Employee rights. In: Hoffman M W, Fredrick R E and Schwartz M S (ed). *Business ethics – readings and cases in corporate morality*. Duquque: McGraw-Hill Higher Education, 2001.
21. Bowie N E. A Kantian theory of meaningful work. *Journal of Business Ethics* 1998; 17 (9–10): 1083–1092.
22. Nord E. *Cost–value analysis in health care. Making sense out of QALYs*. Cambridge: Cambridge University Press, 1999.
23. Rawls J. *A theory of justice*. Oxford: Oxford University Press, 1972.
24. Johannesson M, Hansson S O, Rudén C and Wingborg M. Risk management – the Swedish way(s). *Journal of Environmental Management* 1999; 57 (4): 267–281.
25. Wirth D A and Silbergeld E K. Risky reform. *Columbia Law Review* 1995; 95 (7): 1857–1895.
26. Health and Safety Commission. *A strategy for workplace health and safety in Great Britain to 2010 and beyond*. London: HSC, 2004. www.hse.gov.uk/aboutus/hsc/strategy.htm.
27. Weil D. Enforcing OSHA: the role of labor unions. *Industrial Relations* 1991; 30: 20–36.
28. Weil D. Building safety: the role of construction unions in the enforcement of OSHA. *Journal of Labor Research* 1992; 12: 121–132.
29. Weil D. If OSHA is so bad, why is compliance so good? *RAND Journal of Economics* 1996; 27 (3): 618–640.
30. Halgunset J and Svarva A. *Arbeidstilsynet – politi eller rådgiver*. Trondheim: Institutt for Industriell Miljøforskning, 1980: 125–126.

31. Johnstone R. Improving worker safety: reflections on the legal regulation of OHS in the 20th century. *Journal of Occupational Health and Safety – Australia and New Zealand* 1999; 15: 521–526.
32. Gray W B and Scholz J T. Analyzing the equity and efficiency of OSHA enforcement. *Law and Policy* 1991; 13: 185–214.
33. Gunningham N. Integrating management systems and occupational health and safety regulation. *Journal of Law and Society* 1999; 26 (2): 192–214.
34. Gray W B and Jones C A. (1991a) Are OSHA health inspections effective? A longitudinal study in the manufacturing area. *Review of Economics and Statistics* 1991; 73: 504–508.
35. Gray W B and Jones C A. (1991b) Longitudinal patterns of compliance with Occupational Safety and Health Administration health and safety regulations in the manufacturing sector. *Journal of Human Resources* 1991; 26: 622–653.
36. Hawkins K. *Law as last resort: prosecution decision-making in a regulatory agency*. Oxford: Oxford University Press, 2002.
37. Makkai T and Braithwaite J. The dialectics of corporate deterrence. *Journal of Research in Crime and Delinquency* 1994; 31: 347–373.
38. Wilthagen A C J M. *Het overheidstoezicht op de arbeidsomstandigheden*. Wolters-Noordhoff, Groningen, PhD thesis. Amsterdam: University of Amsterdam, 1993: 268.
39. Levin R. *Utvärdering av makrometoden vid AI/Malmö. Slutrapport*. Stockholm: Arbetsmiljöverket, Rapport 2002: 4.
40. McCaffrey D P. An assessment of OSHA's recent effects on injury rates. *Journal of Human Resources* 1983; 18:131–146.
41. Smith R S. The impact of OSHA inspections on manufacturing injury rates. *Journal of Human Resources* 1979; 14: 145–170.
42. Cooke W and Gautschi F H. OSHA, plant safety programs and injury reduction. *Industrial Relations* 1981; 20: 245–257.
43. Gray W B and Scholz J T. Does regulatory enforcement work? A panel analysis of OSHA enforcement. *Law & Society Review* 1993; 27: 177–213.
44. Scholz J T and Gray W B. OSHA enforcement and workplace injuries: a behavioral approach to risk assessment. *Journal of Risk and Uncertainty* 1990; 3: 283–305.
45. Baggs J, Silverstein B and Foley M. Workplace health and safety regulations: impact of enforcement and consultation on workers' compensation claim rates in Washington State. *American Journal of Industrial Medicine* 2003; 43 (5): 483–494.
46. Magat W A and Viscusi W K. Effectiveness of the EPA's regulatory enforcement: the case of industrial effluent standards. *Journal of Law and Economics* 1990; 33: 331–360; see also Ringquist E J. Is 'effective regulation' always oxymoronic?: the States and ambient air quality. *Social Science Quarterly* 1995; 76: 69–87.
47. van Klaveren M and Tijdens K. Telework: regulated or deregulated. In: Book of Abstracts, 22nd International Labour Process Conference 5–7 April 2004, Amsterdam. Amsterdam Institute for Advanced Labour Studies, University of Amsterdam.
48. Persson A J and Hansson S O. Privacy at work – ethical criteria. *Journal of Business Ethics* 2003; 42 (1):59–70.
49. Council Directive 89/391/EEC of 12 June 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work (1989). Official Journal L 183, 29/06/1989.
50. Walters D (ed). *Regulating health and safety management in the European Union: a study of the dynamics of change*. Bruxelles: PIE-Peter Lang, 2002.

