Environmental Enforcement When 'Inspectability' Is Endogenous: A Model With Overshooting Properties

ANTHONY G. HEYES
Department of Economics, Birkbeck College, University of London, 13–15, Gresse Street, London, W1P 1PA, England

Abstract. If a firm can influence its monitorability vis-à-vis an environmental regulator, it is shown that increasing the thoroughness of inspections induces the firm to substitute towards more transparent technologies, whilst increasing their frequency may cause substitution the other way. Perversely, when the effect of such substitution is taken into account, an increase in the frequency of inspections (or, equally, the stringency of penalties) may worsen the firm's environmental performance. The agency should favour more thorough inspections than existing theory suggests, particularly in sectors where the scope for such substitution is great. Moreover, when monitorability adjusts only sluggishly to policy shocks (because it is an embodied characteristic of capital, for example) the environmental impacts of increased frequency and increased thoroughness well over- and under-shoot their respective long-run impacts. In assessing regulatory reform, therefore, it is important to leave sufficient time for the class of adjustments identified to occur. The possibility of overshooting can be used as an alternative to existing 'regulatory capture' theories to explain why the efficacy of some classes of regulatory reform may fade through time.

Key words. Environment, regulatory enforcement, probability of compliance

1. Introduction

For environmental regulations to protect the environment compliance with those regulations must be fully or partially enforced. This is true whether the regulatory instrument being used is an engineering requirement, an emissions target or some other. Even when a system of tradable pollution permits is being operated, the integrity of the system relies on there being an enforcement agency verifying that agents do not exceed their respective entitlements.

In most contexts the monitoring technology available to environmental enforcement agencies is inaccurate (see, for example, Segerson, 1988; Miceli, 1990), the inaccuracy reflecting both (i) uncertainty regarding the intrinsic properties of the underlying phenomena and (ii) the inadequacy of the monitoring technology (Russell, 1990; Russell *et al.*, 1986). It is also likely to be effected by the characteristics of the polluting technology. This inaccuracy means that even if a firm which is failing to comply with some environmental regulation is inspected, the inspection may fail to detect that non-compliance, or it may detect the non-compliance but fail to yield enough

Environmental and Resource Economics **4**: 479–494, 1994.
© 1994 *Kluwer Academic Publishers. Printed in the Netherlands.*

evidence to permit the regulatory agency to levy a penalty (see, for example, Garvie and Keeler, 1993; Mintz, 1988).

Clearly, the enforcement agency can increase the probability that an inspection of a non-compliant firm will lead to that firm being penalised by increasing the 'thoroughness' with which it conducts its inspections.[1] The novelty of the analysis presented here is the recognition that the regulated *firm* may be able to influence that probability by "investing in uninspectability".

In a general discussion of law enforcement Bebchuk and Kaplow (1993) note that

> . . . individuals are not all equally easy to apprehend. (U)pon reflection, it seems clear that the probability of apprehension will depend not only on enforcement effort but also on particular characteristics of the actor (Bebchuk and Kaplow, 1993: 217).

We contend that this is equally clear in the context of corporations which break regulatory requirements and, furthermore, that Bebchuk and Kaplow's "ease of apprehendability" parameter would be determined endogenously in such a context (see Humphery *et al.* (1993) for an analogous suggestion in the context of the detection of fraud by financial auditors).

There is extensive anecdotal evidence that firms in the United States and elsewhere have found various ways of reducing the "inspectability" (in the sense outlined) of their activities vis-à-vis environmental inspectors.

Linder and McBride (1984: 327) refer to such exercises on the part of the firm as investments in 'concealment activity', and the potential for them is also discussed by Malik (1990) and Friedman (1981). Garvie and Keeler (1993) note, in a general discussion of instrument selection in the context of environmental enforcement, that the

> . . . intent of firms actively to conceal their actions from regulatory knowledge makes the discovery and verification of violations both difficult and expensive.

In a related paper Kambhu (1989: 105) introduces a variable h which is a measure of ". . . the efforts at deception, to hide non-compliance". In his model the non-compliant firm can, *after* it has been caught, erode the severity of the penalty levied by employing lawyers to 'talk down' the seriousness of the non-compliance on which that penalty will ultimately be conditioned. Similarly, Lee (1984) develops a model of pollution taxes in which the tax can be evaded – the firm can hide some part of any pollution that it causes from the responsible regulatory agency.

One can think of a number of ways in which a firm can 'invest' in reducing the tarnsparency of its activities vis-à-vis the agency, and hence reduce its 'inspectability'.

Consider, for example, the possibility of a firm in a high environment – risk industry setting up parts of the plant that are environmentally benign –

what are sometimes referred to as 'sanitised areas'. These are essentially dummies, operationally redundant, established for the benefit of the inspectors. If inspection is seen as a sampling game in which the inspector tries to find a non-compliant part of the plant (the illegally set effluent outlet amongst the twenty properly set ones, for example), then the firm can decrease the probability that he does so by simply increasing the number of sanitised areas. The spending associated with this type of

> . . . attempt to change operations or employ idle capacity in order to "pass" . . . onsite inspections" (Linder and McBride, 1984: 339)

would constitute one type of investment in uninspectability in our model.[2]

In the United States a firm's constitutional right to privacy (under the 4th amendment) means that the principal environmental inspectorates (notably the National Enforcement and Investigations Center (N.E.I.C.) of the United States Environmental Protection Agency (U.S.E.P.A.) are obliged to conduct at least the initial rounds of their inspection from outside the firm's perimeter fence using remote-sensing devices. If the accuracy of such equipment decreases with distance, the firm can invest in uninspectability simply by buying more land – putting greater distance between the source of the pollutant in question and the nearest point from which surveillance can legally be conducted.[3]

Furthermore, firms may spend money fighting legal battles which influence the *institutional* determinants of inspectability. A major chemical corporation in the U.S. funded a long-running case against the U.S.E.P.A.'s use of aerial photography to search for air pollution (searching which it would otherwise have had to do from the plant-boundary). The District and Appeals courts disagreed over whether this constituted unreasonable search and it was left to the US Supreme court, in May 1986, to rule that only standard photographic equipment could be used during aerial inspection, obliging the U.S.E.P.A. to revise its inspection practices along (presumably) less effective lines.[4] Lobbying of pertinent political and judicial bodies by regulated industries, though costly may reduce the probability that litigation of a given 'quality' (i.e. based on a given body of scientific evidence) by an enforcement agency will result in prosecution – such a process has been referred to as "judicial-" or "legislative-capture" (see Yaeger, 1991; Stigler, 1975). Garvie and Keeler (1993: 5–7) give a similar interpretation to their (exogenous) "judicial attitude parameter" – a high value of the parameter reflecting a pattern of legal decisions in enforcement cases in favour of firms).

More generally, the endogeneity of inspectability can be argued to be intrinsic to the fact that the firm faces a choice of technique. There are, in most manufacturing contexts, an array of alternative technologies that could be used to perform the same task. Each technology will be transparent to regulatory inspection to a greater or lesser extent simply because some designs of plant are intrinsically easier to inspect than others. That is, it is reasonable to suppose that inspectability will often be an embodied characteristic of productive capital.

The examples given here are suggestive of the potential endogeneity of inspectability and serve to motivate the central assumption of our analysis. In the model in Section 2 we assume that there is some way in which the representative firm can, at some cost, reduce the probability that non-compliance with an environmental requirement will be uncovered by an inspection of given thoroughness. Recognising the potential endogeneity of inspectability is shown to have a number of interesting implications for policy prescription and assessment.

2. Stylised Model of the Compliance Decision

In pursuing its productive activities a firm may accidentally cause a discharge of some measure of a prohibited substance into the environment, in which case it is deemed to be 'non-compliant'. The probability of this occuring, π, depends upon the firm's environmental effort level, e, where the marginal physical returns to environmental effort are assumed to be everywhere positive but diminishing: $\pi_e(e) < 0$ and $\pi_{ee}(e) > 0$. We follow Cooter and Ullen (1990) and others in assuming, for simplicity, that the magnitude of the discharge is independent of the probability of its occurrence (compliance is "zero-one" or "binary").

With probability μ the firm will be subject to inspection. If a non-compliant firm is inspected the probability that the inspection will uncover the non-compliance and hence lead to prosecution (and to a penalty, f) is denoted p.[5] This probability depends upon t (≥ 0), the thoroughness of inspection and n (≥ 0), the 'investment' of the firm in uninspectability,[6] where the functional relationship $p = p(t, n)$ is assumed to be twice differentiable. The more spent on a representative inspection the greater is the probability, ceteris paribus, that inspection of a non-complier will uncover the non-compliance and yield a prosecution: $p_t > 0$. we also assume diminishing returns to inspection effort, $p_{tt} < 0$, with p asymptotic to some upper bound (possibly unity) as t increases without bound. When n is large the firm has technical or institutional characteristics that render its environmental compliance-status difficult to assess, making it hard for the inspector to collect the evidence necessary to secure a prosecution. We assume that $p_n < 0$, $p_{nn} > 0$ and that p is asymptotic to some lower bound (possibly zero) as n is increased without bound. The cross-partial derivative p_{nt} is assumed to be positive in the neighbourhood of equilibrium, implying that increasing the thoroughness of inspection serves to reduce the marginal physical productivity of expenditure on n. This constitutes an assumption that whilst an increase in n will decrease p (i.e. more expenditure by the firm on uninspectability reduces the conditional probability of prosecution), that decrease will be smaller when inspections are comparatively thorough.[7]

Though we 'black-box' the prosecution technology in our analysis, we believe these to be plausible restrictions to place on it, and we restrict our analysis to that class of prosecution technologies which satisfy them (see Garvie and Keeler (1993) provide analysis of some such specifications of the technology. The restrictions here are also consistent with those made by Bebchuk and Kaplow, 1993).

The firm chooses how much to spend on environmental effort (e) and how much to 'invest' on reducing its inspectability (n). The former will constitute the numeraire, the relative price of the latter being β. The representative firm is assumed to be risk neutral and thus attempts to minimise its total expected compliance costs, C, which are the sum of its expenditures on e and n and expected penalties:

$$Minimise_{e,n}\{e + \beta.n + \pi(e).\mu.p(t, n).f\}. \tag{1}$$

The final composite term is the expected value of penalties in a given period. It is the probability that the firm, having exerted an environmental effort level e, will (i) be non-compliant, (ii) be inspected and (iii) that the inspection will uncover the non-compliance and lead to successful prosecution, multiplied by the size of the fine. Assuming an interior solution to the firm's problem, the first-order conditions associated with that solution, $\{e^*, n^*\}$, are

$$C_e(e^*, n^*) = \pi_e(e^*).\mu.p_n(t, n^*).f + 1 = 0 \tag{2}$$

$$C_n(e^*, n^*) = \pi(e^*).\mu.p_n(t, n^*).f + \beta = 0. \tag{3}$$

These are straightforward to interpret. Increasing e reduces the probability that the firm will be penalised by reducing the likelihood that it will be non-compliant, whilst increasing n reduces that probability by reducing the chance that an inspection will detect such non-compliance as does exist. In each case optimality requires that the marginal saving generated in expected penalties be set equal to the unit cost of the respective choice variable. Increases (decreases) in e will correspond to real improvements (reductions) in environmental quality.

2.1. CONVENTIONAL WISDOM RESTATED: REGULATORY REFORM WHEN INSPECTABILITY IS PREDETERMINED

Before examining the case in which monitoring noise is endogenous, it is useful to consider the impact of policy reform when n is fixed – when the firm is assumed unable, for whatever reason, to affect its inspectability. We take the policy conclusions derived under this circumstance to constitute 'conventional wisdom' (*CW*):

PROPOSITION 1: When n is exogenous e^*, the firms equilibrium level of environmental protection effort is (i) unambiguously increasing in μ, t and f, and (ii) unambiguously decreasing in n.

PROOF: By application of the implicit function theorem to Equation 2 (see Appendix 1).

Thus, *CW* predicts that the firm will spend more effort on ensuring compliance when inspections are more frequent and/or more thorough, and when fines for being shown to be in violation of environmental regulations are larger. These are consistent with the qualitative results of a number of analyses, including Storey and McCabe (1980), and Harford (1987). That de^*/dn is negative is also apparent implying that firms will, other things being equal, exercise less environmental care when the monitoring technology available to the inspectorate is less accurate.

2.2. REGULATORY REFORM WHEN INSPECTABILITY IS ENDOGENOUS: COMPARATIVE STATICS

The comparative static predictions of *CW* turn out, however, to be quantitatively and (in some cases) qualitatively unrobust to the endogenisation of inspectability.

When the firm has discretion over n and e it can reduce the frequency with which it expects to be fined for violations of the environmental regulation by altering either one or both. Increasing e reduces the frequency of violation, increases in n reduce the probability that any particular violation will be detected. Thus whilst the aim of tightening environmental regulation is to encourage the firm to take more precautions to prevent accidental violations (i.e. to increase e) it may have the incidental impact of inducing it to be more or less vigorous in its efforts to obscure and obstruct the enforcement process.

It can be shown that the endogenisation of n serves to reinforce the efficacy of policies aimed at increasing the *thoroughness* of inspection, but to reduce the benefit from increases in the *frequency* of inspections (i.e. the probability that a given firm will be inspected in a given period) or the level of fines.

The comparative static impacts of changes in each parameter on the level of environmental care taken by the representative firm are characterised in Proposition 2.

PROPOSITION 2: When n is endogenous: (i) e^* is unambiguously increasing in t, but its response to changes in μ and f are ambiguous. (ii) n^* is unambiguously decreasing in t, but its response to change in μ and f are ambiguous.

A necessary and sufficient condition for $(de^*/d\mu) > 0$ and $(de^*/df) > 0$ is that $[p_n^2 - p \cdot p_{nn}] < 0$ in the vicinity of equilibrium.[8]

PROOF: By application of Cramer's rule to the system of Equations 2 and 3 (see Appendix 2).

An increase in the thoroughness of the representative inspection unambiguously increases the environmental effort level of the representative firm. Endogenising inspectability serves to reinforce the qualitative prediction of *CW* because increases in t reduce the marginal productivity (in terms of expected penalty reduction) to the firm of expenditure on n, and induce the firm to reduce such expenditures. With n lower, however, non-compliance is more likely to result in prosecution and the firm finds it profitable to increase its efforts to prevent such non-compliance, e. The implication is that the policy of increasing the thoroughness of inspections is even more effective at protecting the environment when n is endogenous than was predicted by *CW*.

The same cannot be said for increases in inspection frequency. If $[p_n^2 - p \cdot p_{nn}] > 0$ in the vicinity of equilibrium then increasing the frequency of inspections or the size of the penalty will worsen the environmental performance of the firm.[9]

As can be seen from the proof of Proposition 2, the endogenous adjustment of n serves to reduce the impact of increases in μ on e^*.

The direct effect of an increase in μ on e works in the 'right' direction: non-compliance is made more likely to be detected and so the firm increases its efforts at ensuring that it does comply (i.e. it increases e). The increase in μ has a secondary effect through changes in n, however, which works in the opposite direction. Because the firm expects to be inspected more frequently, it becomes more valuable to it to have comparatively uninspectable plant – the increase in μ increases the marginal return to the firm of expenditure on n and induces an increase in it. With n higher non-compliance is less likely to be discovered and the firm has an incentive to *reduce* its environmental effort, e. The condition stated in the Proposition is necessary and sufficient to ensure that the direct effect dominates. If it does not hold then the policy of increasing the frequency of inspections causes a net *reduction* in e^*, contrary to *CW*. This discussion also illustrates the intuition behind the relationship outlined in note 9: The firm increases at least one of its levels of effort in response to an increased expected penalty $\mu.f$. The firm may choose to reduce its level of environmental efforts, but only if it also increases its efforts at making detection more difficult.

There is no particular reason to believe, *a priori*, that the condition which rules this out will be satisfied (requiring, as it does, an assumption that the first derivative of $p(t, n)$ with respect to n will be sufficiently small in absolute terms), and even in those cases where e^* *is* increasing in μ the size of the policy-multiplier is reduced by the endogenisation of n. Part if not all of the impact of an increase in μ is "absorbed" by the firm through substitution towards less easily inspected technologies.

Recognising that μ and f enter the firm's minimisation problem symmetrically it is clear that the impact of a change in f will be comparable to that of an increase in μ. The fact that raising the level of fines is something that can be done by fiat and therefore, at least in principle, without increasing enforcement costs, has made it a popular policy possibility (see Storey and McCabe (1980: 242), Stigler (1975); Becker (1968) put the original case for maximal penalties in the context of the 'economics of crime' literature.) The recognition that the condition in Proposition 2 may not be satisfied raises the possibility that increasing the severity of fines faced by firms caught breaking environmental rules could actually induce a reduction in environmental protection effort. In any event, the endogenisation of n reduces the efficacy of penalty hikes as a way of enhancing compliance.

The comparative statics offer another insight into the regulatory dilemma of how far limited enforcement resources should be spent on increasing the frequency of inspections, and how far on increasing the thoroughness with which each inspection is carried out. Our results suggest that, at least in industries where the technological and institutional environment is such that firms have significant discretion over their own inspectability, prudent policy should be biased towards less frequent but more thorough inspections than existing regulatory theory would suggest. Increases in their frequency may even prove counterproductive.

3. The Dynamics of Compliance Adjustment: A Theory of 'Regulatory Fade'

It is interesting to decompose the comparative static results of the last section to examine what implications the endogenisation of n might have for the regulatory adjustment process.

For the purposes of the analysis in this section we assume that the firm can adjust e instantaneously but that n can be adjusted only sluggishly. That e should adjust more rapidly than n seems plausible, given the respective natures of the two variables. Inspectability, an embodied characteristic of industrial plant and its institutional environment, is unlikely to be adjustable except in the medium to long-run (e, if you like, is the effort that the operator takes to use a given technology carefully whilst n is a characteristic of the technology itself). Taking the embodied view of n, the rate at which it will converge on its new equilibrium level in the wake of an unanticipated policy shock will depend upon the rate at which capital is turned-over in the industry in question.[10]

Without developing an explicit model of dynamic adjustment, diagrams can be used to compare the long- and short-run implications of unanticipated changes in the enforcement parameters. Of particular interest is the possibility that the environmental performance of the firm will over- or under-shoot its long-run level in response to policy shocks.

Remark: The environmental effort level may over-shoot or under-shoot its long-run level in response to an unanticipated increase in inspection frequency. The endogenous adjustment of inspectability always serves to diminish the long-run efficacy of the policy.

Consider Fig. 1. The two curves represent the loci of points associated with combinations of e and n which satisfy the respective first-order condition. It is straightforward to confirm that both are downward sloping in $\{n, e\}$-space by implicit differentiation of the pair of first-order conditions associated with the firm's minimisation problem:

$$\frac{\partial e}{\partial n}(C_n = 0) = -\left[\frac{\pi_e \cdot \mu \cdot f \cdot p_n(t, n)}{C_{nn}}\right] < 0 \tag{4}$$

and

$$\frac{\partial e}{\partial n}(C_e = 0) = -\left[\frac{\pi_e \cdot \mu \cdot f \cdot p_n(t, n)}{C_{ee}}\right] < 0 \tag{5}$$

That the $C_e = 0$ loci is the shallower is can be inferred from the second-order conditions. The intersection of these two loci corresponds to the solution to the firm's problem. The initial equilibrium is labelled 0.

In terms of the diagram, the effect of an increase in μ is twofold. It generates an upward shift in the $C_e = 0$ locus (an increase, ceteris paribus, in environmental effort: $(\partial e^*/\partial \mu)|_{n=\bar{n}} = [\pi_e/\pi_{ee} \cdot \mu] > 0$. It also generates a rightward shift in the $C_n = 0$ locus: $(\partial e^*/\partial \mu)|_{e=\bar{e}} = -[p_n/\mu \cdot p_{nn}] > 0$. As was shown in Proposition 2, the new long-run equilibrium (labelled 2), once established, could involve a higher or lower value of both n and e (which can be verified by inspection of Fig. 1).

The case illustrated in Fig. 1 is that in which the equilibrium responses of e and n to the increase in μ are positive (i.e., the comparative static result coincides, qualitatively, with *CW*). The long-run equilibrium moves from 0 to 2. In the short-run, however, n is fixed and thus, at the moment of impact,

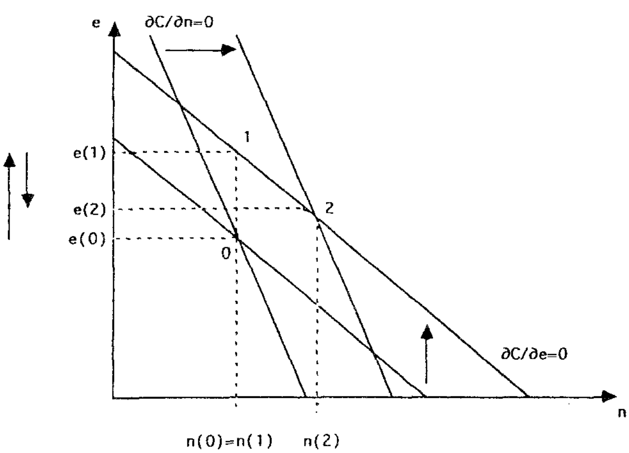

Fig. 1. Effect of an increase in μ.

the firm jumps to a temporary equilibrium at 1. The move from 1 to 2 occurs gradually as n adjusts (sluggishly) to the change in enforcement regime. The time-path for e is only piecewise continuous, at the moment at which the policy reform is instigated e jumps from $e(0)$ to $e(1)$, it then converges on $e(2)$ as n converges on its new equilibrium level.

In contrast, if the $C_n = 0$ locus shifts comparatively little when μ is increased, appropriate alteration of Fig. 1 shows e will undershoot its long-run level: the impact effect is smaller than the steady-state effect.

In both cases it is straightforward to see that the endogenisation of n has reduced the size of the comparative static (i.e. long-run) effect of the change in μ on e. The increase in μ is associated with the rightward shift in the $C_n = 0$ locus. This moves the new equilibrium (labelled 2) downwards along the negatively sloped $C_e = 0$ locus, reducing $e(2)$. The greater is the shift in the $C_n = 0$ locus (i.e. the greater is the adjustment in n) the greater is the extent of this downward movement, *ceteris paribus*. Clearly, if n adjusts sufficiently much then $e(2)$ can come to be less than $e(0)$, and the policy of increasing inspection frequency to improve environmental performance proves counterproductive. This is the case in which the sufficiency condition in Proposition 2 is not satisfied.

As such our model provides an alternative to "regulatory capture" theories for why the efficacy of policy reform may tend to wear-off or 'fade' through time. Regulatory capture refers to the process through which the regulatory system, its institutions or agents come to find themselves 'on the side of' the firm or firms for which they are responsible, and thus more lenient towards wrong-doing (Posner, 1974 and Stigler, 1975 provide the seminal contributions; see also Williamson, 1988). Thus

> . . . regulatory agencies may over time come to defend the interests of the industries for which they are responsible . . . the close ongoing association between the industry and the regulatory agency can increasingly lead to an altered perception of what is desirable in the trade-off between the interests of the industry and the interests of the public at large (Hay and Morris (1991: 631)).

Consider, for example, the policy of increasing μ through increasing the number of inspectors. The initial impact of this on the propensity of the regulated industry to comply with environmental requirements is great, but could be expected to diminish through time as the newly extended agency and its personnel gradually come to be "captured" by the firms they are employed to police. The mechanism through which such capture is supposed to occur is not well specified, but the theory is framed in terms of changes in the characteristics and attitudes of the *regulator* as the enforcement relationship unfolds (Kalt and Zupan, 1984).

The overshooting properties identified here offer an alternative explanation for the phenomenon, with 'regulatory fade' being generated by the adjustment dynamics of the model. Here, in contrast to the regulatory capture theories, it is the characteristics of the *regulatee* which change as the enforcement relationship develops. Consider Fig. 1 again. When μ is increased the firms new equilibrium point becomes point 2, with less inspectable technologies being employed than was the case before the policy reform (i.e. n^* has increased). The nature of n, however, is that it cannot be changed instantaneously. In the short-run then, the firm is stuck with n at its old (and now inappropriately low) pre-policy-shock level. The firm reoptimises with respect to e, given the new value of μ and the constraint that n is predetermined at $n(0)$. This reoptimisation involves a substantial increase in environmental effort, with e jumping to $e(1)$. Not all of this additional exertion will, however, be sustained. In the longer-run the firm reduces the inspectability of its operations (by substitution towards plant with higher embodied values of n, or some other means). As n adjusts towards its new equilibrium level, $n(2)$, e is continuously reoptimised 'en route' to take account of those adjustments. Overall, then, e overshoots its long-run level, implying that the efficacy of the policy reform fades through time, 'as if' the new apparatus of inspection was subject to progressive capture.

The same type of overshooting will characterise increases in the level of penalties, f. The impact effect on environmental quality (proxied by e) will be comparatively large but will be eroded as n gradually adjusts towards its new equilibrium level. As was demonstrated in Proposition 2, this erosion may be so great as to outweigh the initial, environmentally beneficial impact. In that case the long-run impact of an increase in penalties upon environmental quality would be negative, the principal effect of the increased stringency being to encourage the industry to channel more of its efforts into obstructing the enforcement process.

Remark: Given the assumptions introduced in the text, the environmental effort level undershoots its long-run level in response to an unanticipated increase in inspection thoroughness. The endogenous adjustment of monitoring noise serves to enhance the long-run efficacy of the policy.

Consider Fig. 2, where the loci are as before.

An increase in t induces an upward shift in the $C_e = 0$ locus since $(\partial e^*/\partial t)|_{n=\bar{n}} = -[\pi_e \cdot p_t/\pi_{ee} \cdot p] > 0$. It also causes a leftward shift in the $C_n = 0$ locus, as dictated by $(\partial n^*/\partial t)|_{e=\bar{e}} = -[p_{nt}/p_{nn}] < 0$.

As can be seen from the diagram, then, the equilibrium effect of the reform is to increase e and decrease n (confirming part of the result of Proposition 2). Notice that if n adjusts only sluggishly, as we are assuming, e will *under*shoot its long-run level. The impact effect of the policy will be for e to jump from $e(0)$ (its initial equilibrium level) to $e(1)$, from where it will gradually converge on $e(2)$ as n is revised downwards towards its new equilibrium level, $n(2)$. Environmental performance, then, exhibits undershooting – the observed efficacy of the policy reform increasing through time.[11]

Furthermore, it can readily be seen that the endogenous response of n to

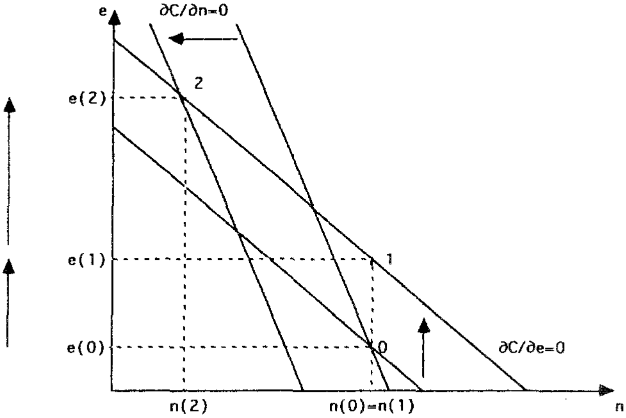

Fig. 2. Effect of an increase in t.

the policy stimulus serves to enhance that policy's efficacy (for the reasons outlined in the discussion of Proposition 2).

4. Conclusions

In a number of industries there exists scope for the regulated firm to obstruct the process whereby environmental regulations are enforced by manipulating the transparency of their operations vis-à-vis external scrutiny. We have argued, here and elsewhere, that different production technologies will be inherently different in how easy or difficult it is for an inspector to detect environmental violations with a particular degree of conviction. Furthermore, the firm is likely to have discretion over other physical or institutional factors, disembodied from the productive capital itself, which will also influence the transparency of its operations.

Our aim was to examine how the endogenisation of 'inspectability' should affect enforcement policy prescription. The implications of our analysis can be divided into the static and the dynamic. Respectively,

(i) Policy should favour *more thorough* (and, possibly, less frequent) inspections than would be dictated by existing analysis, particularly in industries where the scope for manipulation of inspectability is comparatively large. The efficacy (in terms of environmental protection) of increasing inspection frequency is unambiguously reduced by the endogenisation of inspectability, whilst the efficacy of increasing inspection thoroughness is enhanced. This is because, given the assumptions introduced in the text, the policy changes encourage substitution towards less and more transparent technologies respectively. In some cases the adverse substitution effect may be so great that an increase in inspection frequency may actually prove counterproductive, reducing environmental quality. The results derived for increases in inspection frequency apply equally to increases in the level of penalties.

(ii) In assessing the success of changes in enforcement policy, it is important to leave sufficient time for all (or a sufficient portion of) adjustments in n to occur. Inspectability, to the extent that it is an embodied characteristic of capital, is likely to adjust only sluggishly to policy shocks. The immediate impact of a change in inspection frequency is such that environmental quality may *over*shoot its long-run level. The immediate impact of a change in inspection thoroughness, conversely, is such that environmental quality will *under*shoot. In other words, the former policy will come to look worse through time, the latter will come to look better. We labeled the possibility of overshooting in this context 'regulatory fade', and argued that it provided a coherent alternative to the usual 'regulatory capture' explanation of why the efficacy of some classes of regulatory reform may be expected to wear-off through time.

Appendix 1

Consider the pair of first-order conditions associated with an interior solution to the representative firm's minimisation problem.

$$C_e(e^*, n^*) = \pi_e(e^*)\cdot\mu\cdot p(t, n^*)\cdot f + 1 = 0$$

$$C_n(e^*, n^*) = \pi(e^*)\cdot\mu\cdot p_n(t, n^*)\cdot f + \beta = 0$$

Application of the implicit function theorem to the first of these (treating n as exogenous) yields, after simple cancellation:

$$\left(\frac{de^*}{d\mu}\right) = -\left[\frac{\pi_e}{\pi_{ee}\cdot\mu}\right] > 0$$

$$\left(\frac{de^*}{dt}\right) = -\left[\frac{\pi_e\cdot p_t}{\pi_{ee}\cdot p}\right] > 0$$

$$\left(\frac{de^*}{df}\right) = -\left[\frac{\pi_e}{\pi_{ee}\cdot f}\right] > 0$$

$$\left(\frac{de^*}{dn}\right) = -\left[\frac{\pi_e\cdot p_n}{\pi_{ee}\cdot p}\right] > 0$$

The signs of these are reported in Proposition 1.

Appendix 2

Consider the pair of first-order conditions associated with an interior solution to the representative firm's minimisation problem.

$$C_e(e^*, n^*) = \pi_e(e^*)\cdot\mu\cdot p(t, n^*)\cdot f + 1 = 0$$

$$C_n(e^*, n^*) = \pi(e^*)\cdot\mu\cdot p_n(t, n^*)\cdot f + \beta = 0$$

This simple system can be written in matrix form as:

$$\begin{bmatrix} C_{ee} & C_{en} \\ C_{ne} & C_{nn} \end{bmatrix}\begin{bmatrix} \mathrm{d}e^*/\mathrm{d}X \\ \mathrm{d}n^*/\mathrm{d}X \end{bmatrix} = \begin{bmatrix} -C_{eX} \\ -C_{nX} \end{bmatrix}$$

where X is the parameter in question. Denoting the coefficient matrix Δ, its determinant $|\Delta|$ is positive by the second-order conditions for an interior minimum. Application of Cramers rule permits the comparative static results to be derived:

$$\left(\frac{\mathrm{d}e^*}{\mathrm{d}\mu}\right) = \left[\frac{1}{|\Delta|}\right]\cdot[C_{n\mu}\cdot C_{en} - C_{e\mu}\cdot C_{nn}]$$

$$= \left[\frac{1}{|\Delta|}\right]\cdot[\mu\cdot f^2\cdot\pi\cdot\pi_e\cdot[p_n^2 - p_{nn}\cdot p]]$$

$$\left(\frac{\mathrm{d}e^*}{\mathrm{d}t}\right) = \left[\frac{1}{|\Delta|}\right]\cdot[C_{nt}\cdot C_{en} - C_{et}\cdot C_{nn}]$$

$$= \left[\frac{1}{|\Delta|}\right]\cdot[\mu^2\cdot f^2\cdot\pi\cdot\pi_e\cdot[p_{nt}\cdot p_n - p_{nn}\cdot p_t]] > 0$$

$$\left(\frac{\mathrm{d}n^*}{\mathrm{d}\mu}\right) = \left[\frac{1}{|\Delta|}\right]\cdot[C_{e\mu}\cdot C_{ne} - C_{ee}\cdot C_{n\mu}]$$

$$= \left[\frac{1}{|\Delta|}\right]\cdot[\mu\cdot f^2\cdot p\cdot p_n\cdot[\pi_n^2 - \pi\cdot\pi_{ee}]]$$

$$\left(\frac{\mathrm{d}n^*}{\mathrm{d}t}\right) = \left[\frac{1}{|\Delta|}\right]\cdot[C_{et}\cdot C_{ne} - C_{ee}\cdot C_{nt}]$$

$$= \left[\frac{\mu^2\cdot f^2}{|\Delta|}\right]\cdot[p_t\cdot p_n\cdot\pi_e^2 - p\cdot p_{nt}\cdot\pi\cdot\pi_{ee}]] < 0$$

The signs of the derivatives with respect to f are the same as those with respect to μ since the two terms enter the firms problem symmetrically.

The sign of $(\mathrm{d}e^*/\mathrm{d}\mu)$ is ambiguous and, given the assumptions made about derivatives in the text, can be seen to be the opposite of the sign of $[p_n^2 - p_{nn}\cdot p]$.

Acknowledgement

I would like to thank Robert Cairns, Ngo van Long, Seamus Hogan, Dennis Snower, William Watson, Catherine Liston, two referees of this journal and seminar participants at Birkbeck, Queen Mary and Westfield Colleges of the University of London and at McGill University, Montreal for helpful comments. All errors are mine.

Notes

[1] This could involve using more sophisticated monitoring equipment, taking more samples per inspection-site etc., depending upon the context in question (see, for example, Mintz (1988), Yaeger (1991)).

[2] A variation on this can be found when the regulatory régime involves a self-reporting component (see, for example, Harford (1987)). In such cases there is scope for what Harter (1982) terms 'defensive research', whereby firms invest in spiralling amounts of data generation and excessive documentation in an effort to make it more difficult for external scrutineers to distil the pertinent from the spurious.

[3] See Mintz (1988) or Strock (1990) for more detailed discussion of this possibility (with particular reference to enforcement of the US Clean Air Acts).

[4] This case provides, we believe, an unambiguous example of an investment in uninspectability. What motive could the corporation have had for such an 'investment' unless it was believed that (i) there was some probability that they would, at some stage, be out of compliance with the requisite environmental standard and (ii) that getting the rules of search altered would reduce the likelihood that such non-compliance would lead to penalty?

[5] For analytic convenience we abstract from the possibility of wrongful conviction (i.e. the chance that a compliant firm may be adjudged to be non-compliant). This is consistent with the view of inspection as a process of "finding a fault" – the environmental assessors cannot find a flaw which is not there, but may fail to find one which is. An interesting extension to the model would be to allow for the possibility of false positives. If both the false positive and false negative rates were endogenous, and jointly influenced by both actors, this would serve to increase considerably the ambiguity of our results.

[6] This 'investment' may take an indirect form. For instance, a manager may choose a less physically productive technology, at some cost, because it embodies a higher n value. In this case the n implicitly characterised would have to be uncovered by indirect methods.

[7] An example of a functional form which satisfies these restrictions is $p(t, n) = t/n + \ln(t/n)$.

[8] It is worth noting that when μ and/or f is increased at least one of the effort levels e^* and n^* must increase. This follows since $|\Delta| = \mu \cdot f \cdot [\pi_e \cdot \pi \cdot p_{nn} \cdot p - \pi_e^2 \cdot p_n^2] > 0$ (see Appendix 2). Simple algebra gives $(\pi_{ee} \cdot \pi)/\pi_e^2 \cdot (p_{nn} \cdot p)/p_n > 1$. It follows that at least one of these quotients is greater than one which means that $\pi_{ee} \cdot \pi > \pi_e^2$ and/or $p_{nn} \cdot p > p_n^2$. This is sufficient to imply that at least one of $de^*/d\mu$ and $dn^*/d\mu$ is positive. This result is rather intuitive: The firm increases at least one of its effort levels in response to an increased expected penalty $\mu \cdot f$. In this light the first main result is still interesting, but not so surprising. The firm may choose to reduce its environmental effort, but only if it simultaneously increases its efforts of making detection more difficult.

[9] Kambhu (1989) uses a different mechanism to derive a similar result (Proposition 3, Kambhu, 1989: 108). Andreoni (1991) shows that when the probability of conviction depends on the magnitude of the penalty, higher penalties can encourage crime rather than deter it. Shaffer (1990) and Jones (1980) discuss the importance of the second derivative of the fine function with respect to violation size in determining the direction of the effect which different policy reforms have on the equilibrium frequency of violation. By assuming compliance to be "binary" (i.e. zero-one) we have abstracted from these (very interesting) issues – the results derive in this paper are driven by a quite distinct mechanism to any of these other papers.

[10] In contexts in which n is disembodied from capital, the assumption regarding the rate of adjustment of n will sometimes be harder to justify. An alternative rationale might be to see adjustments in n as coming about through institutional reform (as discussed in the text).

[11] Such a phenomena could be referred to as 'regulatory enhancement' (in contrast to 'regulatory fade' whereby the efficacy of the reform is eroded through time).

Bibliography

Bebchuk, L. A. and L. Kaplow (1993), 'Optimal Sanctions and Differences in Individuals' Likelihood of Avoiding Detection', *International Review of Law and Economics* **13**, 217–224.

Cooter, R. and T. Ulen (1990), *Law and Economics*, Scott, Foresman and Company, London, England.

Friedman, D. D. (1981), 'Reflections on Optimal Punishment, or: Should the Rich Pay Higher Fines?', in R. O. Zerbe, ed., *Research in Law and Economics*, Greenwich, NY: JAI Press, vol. 3, 182–205.

Garvie, D. and A. Keeler (1993), 'Incomplete Enforcement with Endogenous Regulatory Choice'. Discussion Paper #873, Institute for Economic Research, Queens University, Kingston, Canada.

Harford, J. D. (1987), 'Self-Reporting of Pollution and the Firms Behaviour Under Imperfectly Enforceable Regulations', *Journal of Environmental Economics and Management* **14**, 293–303.

Hay, D. E. and D. J. Morris (1991), *Industrial Economics and Organisation*, 2nd Edition, Oxford University Press.

Heyes, A. G. (1992), *Technological Transparency and Regulation.* Mimeo, Department of Economics, McGill University, Montreal.

Humphery, C., S. Turley and P. Moizer (1993), 'Protecting Against Detection: The Case of Auditors and Fraud', *Accounting, Auditing and Accountability Journal* **6**(1), 39–62.

Jones, C. A. (1989), 'Standard-Setting with Incomplete Enforcement Revisited', *Journal of Policy Analysis and Management* **8**, 72–87.

Kalt, J. P. and M. A. Zupan (1984), 'Capture and Ideology in the Economic Theory of Politics', *American Economic Review* **74**(3), 279–300.

Kambhu, J. (1989), 'Regulatory Standards, Non-Compliance and Enforcement', *Journal of Regulatory Economics* **1**(2), 103–114.

Lee, D. R. (1984) 'Monitoring and Budget Maximisation in Pollution Control', *Economic Inquiry* **18**, 565–575.

Linder, S. H. and M. E. McBride (1984), 'Enforcement Costs and Regulatory Reform: The Agency and Firm Response', *Journal of Environmental Economics and Management* **11**(2), 327–346.

Malik, M. S. (1990), 'Avoidance, Screening and Optimal Enforcement', *Rand Journal of Economics* **21**(3), 341–353.

Miceli, T. J. (1990), 'Optimal Prosecution of Defendants Whose Guilt is Uncertain', *Journal of Law, Economics and Organisation* **6**(1), 187–201.

Mintz, J. A. (1988), 'Agencies, Congress and Regulatory Enforcement: A Review of EPA's Hazardous Waste Enforcement Effort, 1970–1987', *Environmental Law* **18**, 683–759.

Posner, R. (1974), 'Theories of Economic Regulation', *Bell Journal of Economics* **5**, 335–358.

Russell, C. S., W. J. Vaughan and W. Harrington (1986), *Enforcing Pollution Control Laws*, Johns Hopkins University Press, Baltimore, MD.

Russell, C. S. (1990), 'Monitoring and Enforcement', Chapter 7 in P. R. Portney, ed., *Public Policies for Environmental Protection*, Washington DC: Resources for the Future.

Segerson, K. (1988), 'Uncertainty and Incentives for Nonpoint Pollution Control', *Journal of Environmental Economics and Management* **15**, 87–98.

Shaffer, S. (1990), 'Regulatory Compliance with Non-linear Penalties', *Journal of Regulatory Economics* **2**(1), 99–103.

Stigler, G. (1975), *The Citizen and the State*, Chicago University Press.

Storey, D. J. and P. J. McCabe (1980), 'The Criminal Waste Discharger', *Scottish Journal of Political Economy* **27**, 30–40.

Strock, J. M. (1990), 'Protecting the Environment: EPA After 20 Year', *Trial* (August), 22–27.

Williamson, O. E. (1987), *Antitrust Economics* (1st Edition), Basil Blackwell Ltd., Oxford, England.

Yaeger, P. E. (1991), *The Limits of the Law: The Public Regulation of Private Pollution*, Cambridge University Press, New York, NY.