

OPTIMAL STANDARDS WITH INCOMPLETE ENFORCEMENT

W. KIP VISCUSI AND RICHARD J. ZECKHAUSER

Formal analyses of regulation typically assume that all firms will comply with the regulations imposed. Such analyses frequently examine the effect of a regulation, say a pollution standard, on a "representative" firm. Reality is far more complex, with the responses of firms to a given regulation varying widely, depending on their individual circumstances. In particular, some firms will choose to risk penalties for noncompliance, rather than incur the costs of meeting the standard. As a standard is made more stringent, the performance of firms that comply increases, but additional firms will choose not to comply. Thus, as we demonstrate in this paper, at some point further tightening of the standard may lower overall performance.

I. Introduction

The economics of standard setting has traditionally been approached from two directions. The first assumes full compliance, and seeks to establish an optimal standard by equating the marginal benefits of a raised standard with the marginal costs of achieving it.¹ The second takes the standard as given and investi-

Albert Nichols gave us helpful comments and suggested future lines of research we shall explore jointly with him. Zeckhauser's research was supported by NSF grant SOC(77-16602).

¹ For an early approach along these lines, see the essays in Maass et al. (1962) and analyses such as Calabresi (1970), Nelson and Neumann (1975), Oi (1973), and Settle and Weisbrod (1977). See Dales (1968) for the statement of the economic motivation for alternatives to standards policies.

Copyright © 1979 by the President and Fellows of Harvard College

Published by John Wiley & Sons, Inc.

0033-3646/79/0027-0437\$01.00

gates the optimal structure for penalties to assure the proper degree of compliance, which may be less than total.²

Standard setting should not be divorced from considerations of compliance. The central issue addressed here is how policymakers should select optimal levels of standards when compliance with the standard may not be complete. This approach differs in two major respects from the usual analysis in which a standard is imposed on a representative firm.

First, the standard is not treated as a binding constraint on enterprise actions. Instead it is understood that individuals or organizations may find it desirable to violate a standard when the costs of meeting it are high relative to expected penalties if it is not met. Although this phenomenon has not been incorporated in a general analysis of regulatory standards, it has not gone unnoticed. Analysis of criminal activities have long recognized that individual decisions to obey the law will depend on expected criminal sanctions.³ Analyses of the occupational health and safety standards imposed by the Occupational Safety and Health Administration (OSHA) similarly note that enterprise compliance will depend on the cost of meeting the standard and the expected penalties associated with noncompliance; indeed, only about 20 percent of all covered enterprises are found in compliance with OSHA standards.⁴ Within the context of accident law, the existence of a negligence system and due-care standards associated with it may alter individuals' accident-avoiding activities, but does not guarantee that all individuals will undertake at least the prescribed minimal level of care.⁵

The second and more distinctive aspect of the analysis presented here is that the model does not rely on the usual device of examining the response of a representative firm. Instead, it focuses on the characteristics and behavior of the entire regulated popula-

² The literature stressing the selection of the optimal penalty and enforcement strategy includes Becker (1974), Oi (1975), Polinsky and Shavell (1978), Singh (1973), Stigler (1970), and Srinivasan (1973).

³ See, for example, the studies cited in footnote 2.

⁴ For analyses of OSHA, see Oi (1975), Smith (1976), Viscusi (1979), and Zeckhauser and Nichols (1976).

⁵ See especially the work of Diamond (1974) for an analysis of accidents.

tion. Different enterprises are likely to respond in quite different ways to regulation. Some may choose to do nothing, either because they are already in compliance or because compliance is too costly. Others may make major investments to meet the standard, while still others may make only a modest effort to move closer to it. The different natures and degrees of these actions as well as the relative numbers of each type of firm play an integral role in determining the characteristics of the optimal policy.

Although the subsequent analysis will be developed in terms of specific policy problems, such as compliance with product quality standards, the results extend to a variety of regulatory situations in which incomplete compliance and heterogeneous responses are significant. These include health and safety standards imposed by OSHA, fiduciary standards for pension plans stipulated by the 1974 Employee Retirement Income Security Act (ERISA), affirmative-action goals for federal contractors, and accreditation standards for schools and hospitals. Standards imposed on individuals rather than firms, such as the due-care conditions in accident law, can also be analyzed by generalizing the models presented here.

Section II presents a broad view of the regulatory problem, focusing on the motivations for differential responses to standards. The importance of the distribution of population characteristics to the standard-setting process is the subject of Section III. There we show that it may be desirable to employ different standards for different classes of firms, say old and new, or large and small. The insights derived from this framework can be applied to a variety of policy issues such as the optimal size and allocation of an enforcement budget. Such considerations are to be the subject of a subsequent analysis that allows a broader range of considerations, including most particularly cost of compliance, to enter the objective function of the regulatory agency, and enriches the array of regulatory rules the agency might employ.⁶ Section IV presents our conclusions.

⁶ In a preliminary version of this paper, we also applied the model to other issues, such as the choice of the optimal level of fines and the effect of regulation on the existence of firms.

II. Characteristics of the Problem

When the government chooses to impose standards on economic behavior, some form of market imperfection is usually cited as the motivation for intervention. Failure of enterprises and individuals to take into account externalities associated with their actions has led to a variety of pollution-control regulations. External costs of job injuries, such as workmen's compensation benefits, as well as the altruistic concerns of society at large, are among the motivations for regulation of occupational hazards. Similar altruistic concerns and arguments for compensatory justice combine with a belief that the labor markets function imperfectly to provide a potential justification for affirmative-action requirements. The imperfect transmission of information is the primary analytic justification for the regulation of pacemakers, job risks, nuclear safety, pension funding, bank-reserve ratios, the characteristics of institutions such as colleges and hospitals, and indeed for the regulation of product quality in general.

The analysis presented here will not delve into the appropriateness of these regulatory actions or the efficacy of the particular modes of intervention that have been selected. For example, we will not consider the relative attractiveness of standards and taxes or other incentive mechanisms.⁷ Rather we will focus on the narrower problem of determining the optimal level for a standard, given that the standards approach will be used.

Suppose the government imposes a standard on some choice variable k for the enterprise, where for concreteness we will refer to k as the quality level chosen by the firm. The enterprise's response can be illustrated using the simple model sketched in Figure 1. Prior to the imposition of the standard, the firm's payoff (or expected profit) may be expressed as a function of the level of quality k ; it is represented by the curve ABC . The optimizing firm selects the value of k where the curve is highest; in this example that occurs at point B , where the quality level is k_0 . For the situation illustrated, the enterprise's optimum in the absence of regulation is a quality level above the physical minimum, as might be the case with product quality or the financing of a pension

⁷The pollution control literature includes a sizable discussion along these lines. See, for example, Dales (1968) and Settle and Weisbrod (1977).

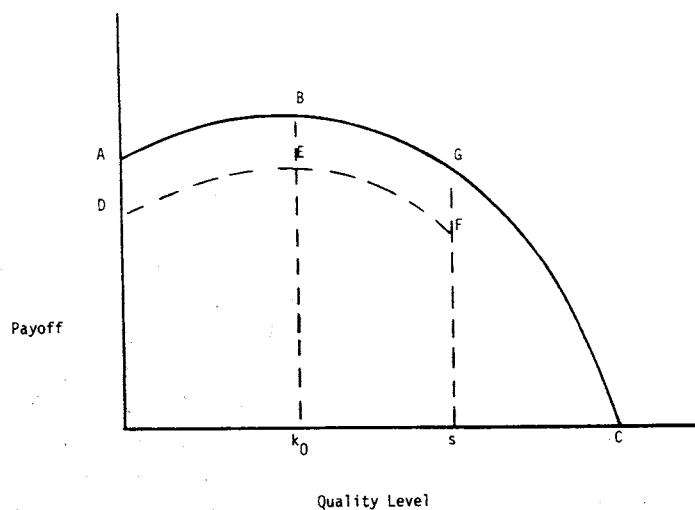
fund. In other instances, such as the control of air or water pollution, the shape of the payoff function may be such that the unregulated firm would find it optimal to undertake no quality-enhancing actions.

Suppose that a minimal quality standard of s is imposed. Firms choosing a quality level below this standard, i.e., those choosing not to comply, have a lower expected payoff than before. Three possible costs are associated with noncompliance. First, the firm may incur fixed costs from being found in violation of the standard. These costs include legal fees, ridicule in the community, a decrease in sales of automobiles subject to recalls, or an increase in wages at a firm found in violation of OSHA standards. If there is no widespread consensus regarding the appropriateness of the standard, this type of cost is likely to be small. Second, the firm risks an expected financial penalty for violating the standard due to fines. Third, if the firm is found to be in violation, it will be forced to incur the costs necessary to comply with the standard, or even to terminate operations altogether. It should be noted that the total cost may be greater if its initial decision is unsatisfactory; for example, a partially effective pollution-control system may have to be replaced altogether in order to meet more stringent standards.

The payoff for quality levels already meeting the standard (the GC portion of ABC) is unaffected by the regulation;⁸ the portion of the representative firm's payoff curve below the standard shifts downward from ABG to DEF . In Figure 1 we show a constant shift. Some of the expected costs of noncompliance may, however, increase the farther the firm's quality level is from the standard. Damage to reputation, for example, is likely to be greater if a firm grossly violates a standard rather than merely falling a bit short. Fines may be scaled to the magnitude of the violation. Finally, the probability of detecting a violation may increase with the extent of noncompliance. The presence of any of these factors would tilt the DEF section of the curve in a counterclockwise direction. In the examples that follow, we assume that there is a constant cost

⁸ This assumption can be modified. Suppose, for example, that firms in the industry experience beneficial group externalities from the imposition of the standard, as might be the case with hospital accreditation. Then the post-standard curve $DEFGH$ would be shifted vertically.

Figure 1. OPTIMAL QUALITY DECISION WITH AND WITHOUT STANDARDS.

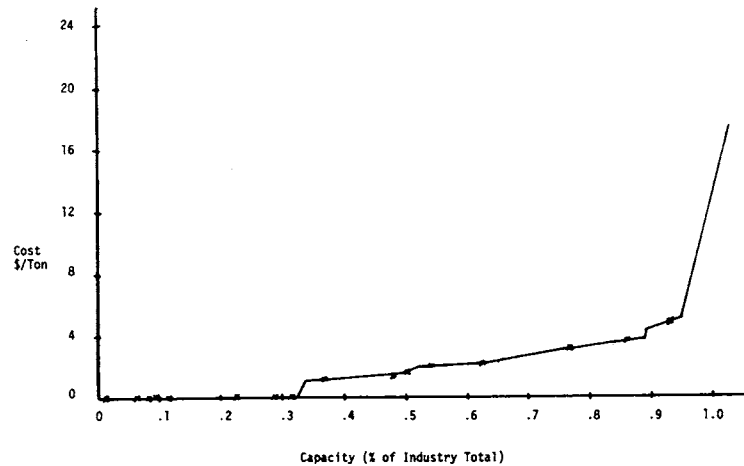


for noncompliance independent of quality level. Relaxing this assumption would not change any of our primary conclusions.⁹

The optimal decision of the firm after imposition of the standard depends on whether E is above or below G . Greater expected penalties lower the value of E and make compliance more attractive. Higher levels of the standard shift G to the southeast and make compliance less attractive. If the expected cost of being in noncompliance is independent of the extent of the violation, increasing the level of the standard simply extends the DEF portion of the payoff curve to the right, where the curve lies below the original payoff schedule by an amount equal to the expected cost of noncompliance. The maximum level of the standard that the firm will adhere to is the value of k for which G represents the same payoff as does E , i.e., where the vertical distance between B and E is the same as that between B and G .

⁹The DEF curve includes the costs of coming into compliance, should a firm be forced to do so, only to the extent that those costs rise if not undertaken initially, or if there is a differential in costs should lower quality levels make detection (hence, ultimate compliance) more likely. If costs were constant over time and probabilities of detection of a violation were constant over quality levels, then the firm should always select the optimal level for them once, knowing that there is some probability (over which it has no influence) that it will ultimately be forced to meet the standard.

Figure 2. DISTRIBUTION OF WATER POLLUTION CONTROL EXPENDITURES IN THE TISSUE PAPER INDUSTRY.



Source: Figure 5 in Leone and Jackson (1978).

Even within a particular industry, the behavior of the different firms may not be identical to that of the representative firm illustrated in Figure 1. Differences in the initial levels of quality and the firms' technological characteristics will result in different compliance costs, with fundamental implications for enterprise behavior. In particular, it seems likely that firms with lower initial quality levels will incur greater costs in meeting a given standard.

Figure 2 illustrates the extent of heterogeneity in compliance for the tissue portion of the U.S. pulp and paper industry.¹⁰ The cost per ton of meeting the 1972 Water Pollution Control Act standards is measured along the vertical axis; it is assumed that each of the 64 mills in the sample adopts the best practicable technology. The cumulative percent of the industry that can meet the standards at a particular cost is measured along the horizontal axis. About one-third of the industry is already in compliance with the standards and will incur no additional expenses. Above that high quality segment of complying firms, costs rise roughly linearly until over 90 percent compliance is achieved. At this upper tail of the distribution, compliance costs become particularly large.

¹⁰ This discussion is based on the analysis in Leone and Jackson (1978).

Figure 3. THE IMPACT OF STANDARDS ON A DISTRIBUTION OF FIRMS.

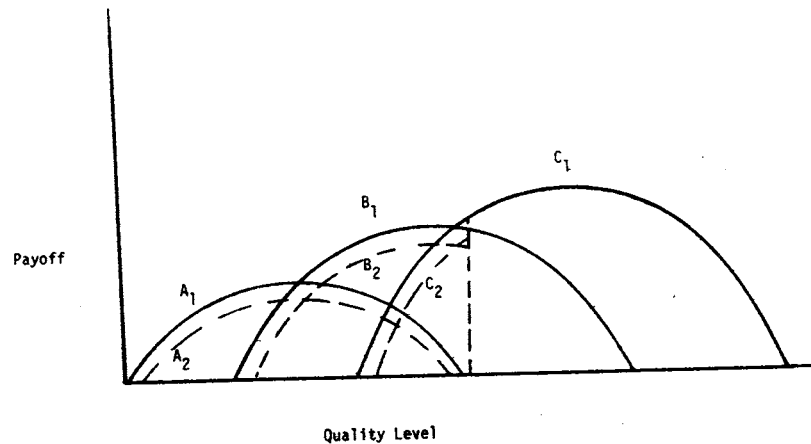


Figure 3 illustrates how regulation will affect firms differently when compliance costs vary in such a manner. Consider three firms in the industry. Suppose the standard s lowers the payoff functions for the three firms from A_1 , B_1 , and C_1 to A_2 , B_2 , and C_2 , respectively, for quality levels not in compliance with the standard. Firm A chooses not to comply, Firm B chooses to raise its quality level to meet the standard, and Firm C is unaffected since it was already in compliance.

These diagrams serve mainly to illustrate the underlying behavior to be embodied in subsequent models. Standards influence the payoffs associated with different quality choices and consequently affect the behavior of the firms. The exact nature of the outcome in terms of the distribution of quality levels achieved hinges on the shape of these payoff curves, the level of the standard, and the distribution of enterprise characteristics. There is no general relationship between the stringency of the standard and the degree of improvement in quality provided; thus raising a standard may either raise or lower average quality.

The nature of the firm's response to different policies can be investigated more fully with the following model: The enterprise has a quality level k , where this pre-regulation level varies among firms in a manner that reflects their particular cost conditions. The cost

c of meeting a quality level standard depends both on the firm's initial level of quality and the level of the standard s , i.e.,

$$c = c(k, s),$$

where

$$\frac{\partial c}{\partial s} \geq 0 \text{ and } \frac{\partial c}{\partial k} \leq 0 .$$

It should be understood that the $c(k, s)$ function will vary from firm to firm. The firm's response to regulation depends on whether compliance costs exceed the expected costs of not meeting the standard. The firm has a probability p of being caught (inspected and forced to comply) and a complementary probability $1-p$ of not being caught. For simplicity it is assumed that all inspections ascertain the firm's quality level with complete accuracy.¹¹ Firms found not complying must pay fine F and incur the costs necessary to meet the standard.¹² A firm will choose to comply with the standard if

$$c(s, k) \leq (1-p) 0 + p [F + c(s, k)],$$

or

$$c(s, k) \leq \left(\frac{p}{1-p} \right) F . \quad (1)$$

¹¹In reality, there is some probability q that the inspector will not detect a violation, in which case the probability of detection p' is simply the probability of inspection p times the probability of detecting a violation during the inspection, or $p' = p(1-q)$. Moreover, p may depend on the extent of the firm's violation; such complicating influences will not be considered here.

¹²If costs differ among firms, as assumed, and if costs of compliance as well as quality level enter the objective function of the regulatory agency, then even firms detected as violating a standard should not automatically be forced to comply. A continuing fine, equal to the value to society of correcting the violation, would be optimal. It would serve as a form of effluent charge, and would permit firms whose costs of compliance are greater than the benefits such compliance would confer to continue to violate. This argument leaves aside the importance of having rules obeyed, tendencies toward the general breakdown of regulatory function, and the like.

The expense the firm is willing to incur to meet the standard varies predictably with the parameters of the problem. As the probability of inspection or being fined increases, the expected penalty is raised so that it is more likely that the cost of compliance will be lower than the expected cost of noncompliance. Higher levels of the standard and lower initial quality levels make it less likely that the firm will find it optimal to comply with the standard since the costs of compliance have been raised relative to the cost of violations. The firms complying with the standard will be those for which k is not below s initially or whose cost of compliance satisfied inequalities [eq. (1)].¹³ The qualitative nature of these results will be incorporated into the optimal intervention models in subsequent sections.

III. Optimal Standards with Heterogeneous Respondents

The natural inclination of policymakers wishing to raise the quality of a product, reduce the presence of carcinogenic substances in the workplace, or diminish pollution, is to set a rigid standard. The reasoning is that stricter standards will promote a higher quality level, assuming that the product remains on the market and the firm remains in business. Within the context of conventional models in which standards act as effective constraints on enterprise decisions, this conclusion is correct.

The actual standard-setting process and the manner in which it generates quality levels, however, involve a critical tradeoff of the type that was illustrated in Figure 3. While higher standards will

¹³The use of eq. (1) to determine the set of firms that will make investments to comply with the standard is straightforward. Suppose that the cost function includes a fixed cost component β and a marginal compliance cost term α which is multiplied by the extent of safety improvement required to compute the variable compliance cost. Then eq. (1) takes the form

$$\alpha(s-k) + \beta \leq \frac{1}{(1-p)} F ,$$

or

$$k \geq s - \frac{\frac{1}{1-p} F - \beta}{\alpha} .$$

The likelihood that a firm will undertake the necessary investment increases with the values of k , p , and F , and decreases with the values of α , β , and s .

result in improved quality provided by enterprises that choose to comply, the percentage of enterprises in compliance will diminish as regulation becomes more stringent.¹⁴ The analysis in this section will illustrate how this phenomenon influences the optimal regulatory strategy. In order to isolate the main features of interest, the enforcement policy will be taken as given externally.

The government's objective is to select the quality standard *s* that will maximize the expected level of quality. This is so commonly the stated objective of a regulatory or other administrative agency that we may neglect to question its reasonableness. Yet if overall social welfare were the goal, the agency would consider the costs of achieving quality as well as quality itself in its objective function.

An agency might employ any of three rationales for focusing on a single objective: (1) Quality levels are far below the optimum, and even after standards are set they will remain so. Thus, any increase in quality will be beneficial on net.¹⁵ (2) The government itself works best on a pluralistic basis. Narrow mission agencies, such as the Consumer Product Safety Commission and the Council on Wage and Price Stability, should compete with one another. From time to time the President, the Congress, the Courts, and public opinion will step in to resolve or strike compromises on issues. (3) The agency's statutory mandate as laid down by Congress requires it to focus on a single objective. (OSHA has frequently set forth this argument in justifying standards that did not consider tradeoffs between benefits and costs.)

Our justification in focusing on the single objective quality is threefold: (a) Since that is frequently what agencies attempt to do, it is worthwhile exploring how they might better perform their tasks.¹⁶ (b) By focusing on the single objective "quality,"

¹⁴More stringent standards also have other ramifications. For example, they may affect the perceived legitimacy of the regulation.

¹⁵This may be true so long as the cost of achieving a given quality level is independent of the way it is achieved, which may not be the case if different firms have different cost curves. Indeed, that is the principal motivation for the argument that economic incentives rather than command and control techniques such as standards should be employed to promote various social purposes. See Schultze (1977) for a wide-ranging (and widely read) discussion of these issues. Progress away from the standards and towards the incentives approach in the real world is hardly so swift as to reduce the urgency of finding improved modes for setting and enforcing standards.

¹⁶This is a particular form of normative approach, well justified when dealing with individual consumers, though frequently scorned when a government agency is the

our central points are made clearly. Subsequent analyses can deal with more complicated objective functions; the central points will still hold and will point in the same direction.¹⁷ (c) Some formulations of quality may, in effect, incorporate cost considerations. For example, different standards may be set for different classes of firms in recognition of cost differentials. Or the penalty structure for noncompliance may be structured to take account of cost. Thus, good faith efforts, i.e., the expenditure of a reasonable level of resources, may be taken into account in mitigating fines; demonstrating excessive compliance costs may be made grounds for exemption from a standard.

We assume that the government has some means of aggregating quality across firms, such as expected number of accidents or average output of pollution per firm. The expected quality level for the economy is the total across firms of their expected quality levels since the expectation of the sum of quality levels is simply the sum of the expectations.

Unlike other analyses of standards, the population affected by the regulation is not assumed to be a single representative firm. As Figure 2 illustrated, even within a narrowly defined industry firms may differ considerably in their cost of meeting the standard. Regulations affecting different industry groups may be subject to far greater differences in quality and compliance costs. For example, the average annual rate of job injuries in 1975 per 100 full-time employees ranged from 0.4 for nondeposit trust companies to 33.7 for special product sawmills.¹⁸

decision maker. We are taking preferences as given, in this case, of the hypothetical regulatory agency, and are seeking to maximize given that set of preferences: When a collection of individuals or a representative body is making a choice, economists like to investigate whether its preferences are well formulated, and in particular whether they represent the true interests of the group. In comparison with many other observers and critics, economists are more skeptical about accepting expressed preferences of groups, but less skeptical about accepting those of individuals.

¹⁷Indeed, mistakenly excluding cost as a consideration biases the choice of standard upward, as does failing to consider compliance. Thus, in the second-best world where an agency follows the single objective "quality," taking account of compliance will push standards in the appropriate direction.

¹⁸These data are from the U.S. Department of Labor (1977). There have been numerous recent discussions of the differential impact of mileage and safety standards on firms within the auto industry. Such standards are generally reported to increase the relative advantage of General Motors over its weaker competitors.

THE APPROPRIATE STANDARD IF IT MUST BE UNIFORM

This heterogeneity of firms has important implications for policy design. Let the initial distribution of quality levels of an item be given by the probability density function $g(k)$. To generate $g(k)$ each firm's initial quality level is weighted by its output.¹⁹ The quality level k is arbitrarily scaled to be in the interval $[0,1]$, where 1 represents the highest attainable quality; if the quality dimension were risk, $k=1$ would represent a riskless product.

The quality of this item is now to be regulated.²⁰ For simplicity, suppose that for every quality standard s there is some critical response point in the population; firms with quality levels below this response point do not comply with the standard, while for those with quality levels above compliance is complete.²¹ In particular, firms with initial quality levels below a critical level $k=h(s)$, where $\partial h/\partial s > 0$, will choose not to comply with the standard because the costs of compliance exceed expected costs of noncompliance. Enterprises with initial quality levels between $h(s)$ and s will undertake the necessary investment to meet the standard, while firms with initial levels equal to or above s will not alter their actions.²²

The government's objective is to maximize the sum of the expected quality levels for the firms' outputs²³ or to select s to maximize

¹⁹ In an earlier version of this paper, we also considered the determination of the optimal fines and the impact of the standards on the affected population, since some firms may go out of business or alter their product mix. The principal alteration is that the first-order conditions must be modified to take into account efficiency losses from the part of the regulated population that is no longer in existence.

²⁰ Analysis of multiple standards for different concerns is for the most part similar to the single-standard approach here. The most important difference for policy is that there may be important externalities among standards. For example, if some standards, such as regulation of toilet seat shape and exit sign placement, are viewed as ridiculous, compliance with other standards may be reduced.

²¹ In a fully general analysis, there will be some fraction of firms at each level of k that choose compliance; this fraction need not equal zero or one.

²² If the cost functions differ in several dimensions, then a multivariate generalization of h can be used so that h will describe a function given the parameter s rather than a discrete response point. If cost functions are multivariate, initial quality level will not be a reliable index on whether a firm will or will not comply.

²³ Note that this formulation includes the rather general set of objective functions of the form

$$\int_0^1 v(q,s)f(q)dq,$$

$$V_1 = \int_0^{h(s)} kg(k)dk + \int_{h(s)}^s sg(k)dk + \int_s^1 kg(k)dk .$$

expected quality level	contribution from noncompliers	contribution from those meeting the standard	contribution from those exceeding the standard
------------------------------	--------------------------------------	---	---

(For the moment we will leave aside complications that arise when noncompliers are inspected, found to be in violation, and then forced to comply. When we return to this issue, we will find that it affects only the magnitude, not the qualitative nature of our results.)

Using Leibnitz's formula, the first-order condition for a maximum of this integral equation is found to be

$$\frac{dV_1}{ds} = 0 = h(s)g(h(s))\frac{dh}{ds} + \int_{h(s)}^s g(k)dk + sg(s) - sg(h(s))\frac{dh}{ds} - sg(s).$$

Rearranging terms, this condition becomes

$$g(h(s)) [s-h(s)] \frac{dh}{ds} = \int_{h(s)}^s g(k)dk. \quad (2)$$

As the standard is raised, additional firms choose not to comply. This loss is represented by the left-hand term in eq. (2). The first two determinants of the loss are the marginal increase in the critical response point due to the increased standard (i.e., dh/ds) and the density function of the items affected by noncom-

where $f(q)$ is the distribution of achieved quality levels for output and $v(q,s)$ is the government's payoff function for output at quality level q when the standard is set at s . Since $v(q,s)$ is in essence a utility function, k is in effect a utility value. Here $f(q)$ is a simple transformation of $g(k)$ in the upper and lower ranges (among those already in compliance or noncompliers); $f(s)$ is the fraction of firms that chooses to raise their quality level to the standard.

²⁴In applying this equation to policy design, one should be careful that one is not at a local optimum rather than a global optimum. This complication is likely to be of some practical consequence since V_1 differs from the usual well-behaved functions subjected to optimization techniques. In particular, the concentration of workers at different levels of safety may be very uneven so that $g(h(s))$ will be such that eq. (2) is satisfied for several values of s .

pliance at the critical response point [i.e., $g(h(s))$]. The product of these two terms represents the magnitude of the population affected by additional noncompliance. The final component of the left-hand term represents the magnitude of their quality loss [i.e., $s - h(s)$] from additional firms that choose not to comply with the standard. The right-hand term in eq. (2) is the number of items produced by firms undertaking investments to meet the standard and is a measure of the gain from raising the standard, since the extent of the benefit equals the marginal quality improvement (1 unit) multiplied by the quantity of items. As the first-order condition indicates, the distribution of quality levels in the market and the behavior and concentration of enterprises at the critical response point are of central importance in selecting an optimal standards policy.

The most important general principle generated by these results is that the optimal standard with incomplete compliance typically will be below, and never will be above, the level of the optimal standard with full compliance. If total compliance were assured the expected quality level would be maximized by setting s equal to 1. With incomplete compliance, however, the optimal standard will be given by an interior point satisfying eq. (2), except in the rare case of a corner solution. At or above the optimal level of s , raising the standard will reduce quality since more stringent standards will not lead to sufficient improvements to offset the losses from firms foregoing investments to improve quality. In particular, the marginal cost term represented by the left-hand side of eq. (2) will exceed the marginal benefits term on the other side of the equation.

COMPLIANCE AFTER INSPECTION

In addition to the anticipatory effect of regulation deriving from the expected penalties for noncompliance, there is a further quality effect stemming from actions taken by firms that are inspected. Suppose that there is some probability p that a firm choosing noncompliance will be inspected and that the penalty structure has been designed to assure that a firm found in violation of the regulation brings its quality level up to the standard. Then the expected quality V_1 equation becomes

for fertilizer spreaders. A well-designed regulatory policy will take this heterogeneity into account, recognizing that different firms have different compliance costs, that the distribution of quality varies by product and industry, and that consumers and workers may have different tradeoffs between quality and money.

If there were no difficulties such as greater coordination problems associated with non-uniform standards, it would be optimal to permit standards to vary for each firm for each product and, perhaps most narrowly, for each unit consumed by individuals. The advantage of completely variable standards is analogous to the advantage of price discrimination as against uniform pricing for a monopolist.

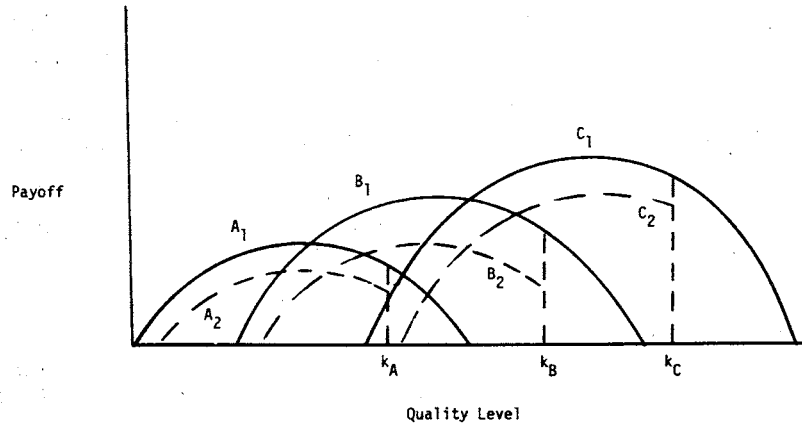
Consider the situation in which the objective is to maximize average quality. Let the optimal uniform standard be k_2 for the population of firms illustrated in Figure 2. Since neither Firm *A* nor Firm *C* alters its actions in response to the regulation, the efficacy of the policy would clearly be enhanced if quality at these two firms were improved as well as quality at *B*. The promulgation of standards k_A , k_B , and k_C for the three respective firms will result in higher quality levels at each of the three firms. As illustrated in Figure 4, each firm finds it optimal to choose a quality level equal to the firm-specific quality standard. The analogues of the various equations describing the optimal regulatory policies can be modified similarly by simply applying the earlier results to each subpopulation that is considered.²⁵

Although completely flexible standards will undermine the principal advantage of standards policies—their simplicity—some distinctions can be profitably drawn. Different safety standards for old facilities and new investments, different certification requirements for previously practicing psychologists and new entrants, and other variations for actions that involve sunk costs as opposed to new allocations appear to be obvious candidates for a discriminatory policy. In particular, the critical response point for new activities is likely to be higher and more responsive to regulatory action.

Another dimension along which distinctions appear most important is the character of the market in the particular industry. If the

²⁵To do this, add subscripts to all of the variables to indicate the population to which the standards apply.

Figure 4. THE IMPACT OF NON-UNIFORM STANDARDS.



Consumer Product Safety Commission imposes standards of the same quality on all products or if OSHA is equally demanding in its regulation of risks, the policy is unlikely to be as successful as a more flexible approach.

IV. Conclusion

Consideration of the differential response of firms to regulatory standards and the corresponding distribution of firm characteristics introduce new complexities that cannot be captured by generalizing the analysis of a representative firm. Unlike usual standards models, the differential response frameworks demonstrate that higher levels of standards may diminish the quality levels provided since more firms will choose to violate the standard when the costs of compliance become too great. Future analyses will examine modifications that result in these findings when the regulatory policy takes account of firms' compliance costs as well as of average quality achieved.

The principal concern of those designing the regulatory policy should not be the response of the average firm in the market, but instead should be the reaction and concentration of firms at the critical response point in the affected population. To facilitate policy design along these lines, the information provided to policy-

makers will also have to be different. Instead of focusing on average injury and accident rates, for example, analysts will require information on the entire distribution of outcomes.

Analysis of the differential responses also suggests that the use of uniform standards may not be optimal. The conventional argument against rigid standards is that they induce rampant inefficiencies since they ignore differences in compliance costs. Even if one ignores these efficiency losses, however, rigid standards are not well-suited to the promotion of the desired quality improvements. Non-uniform standards offer much greater potential for effectiveness and are particularly appropriate when old and new investment activities entail different compliance costs.

BIBLIOGRAPHY

- Becker, Gary: "Crime and Punishment: An Economic Approach," in G. Becker and W. Landes (Eds.), *Essays in the Economics of Crime and Punishment*, Columbia University Press, New York, 1974, pp. 1-54.
- Calabresi, Guido: *The Costs of Accidents*, Yale University Press, New Haven, 1970.
- Dales, J.H.: *Pollution, Property, and Prices*, University of Toronto Press, Toronto, 1968.
- Diamond, Peter: "Accident Law and Resource Allocation," *Bell Journal of Economics and Management Science*, 5(2):366-405, Autumn 1974.
- Leone, Robert, and John Jackson: "The Political Economy of Federal Regulatory Activity," Working Paper HBS 79-6, Harvard Business School, 1978.
- Maass, Arthur, et al.: *Design of Water-Resources Systems*, Harvard University Press, Cambridge, 1962.
- Nelson, Jon P., and George R. Neumann: "Labor Productivity and the Coal Mine Health and Safety Act of 1969," Pennsylvania State University, processed 1975.
- Oi, Walter: "The Economics of Product Safety," *Bell Journal of Economics and Management Science*, 4(1):3-28, Spring 1973.
- : "On Evaluating the Effectiveness of the OSHA Inspection Program," University of Rochester, processed 1975.

- Polinsky, A. Mitchell, and Steven Shavell: "The Optimal Tradeoff Between the Probability and Magnitude of Fines," HIER Discussion Paper No. 597, 1978; forthcoming in the *American Economic Review*.
- President's Report on Occupational Safety and Health*, U.S. Government Printing Office, Washington, D.C., 1976.
- Schultze, Charles: *The Public Use of Private Interest*, Brookings Institution, Washington, D.C., 1977.
- Settle, Russell, and Burton Weisbrod: "Governmentally-Imposed Standards: Some Normative and Positive Aspects," University of Wisconsin, processed 1977.
- Singh, Balbir: "Making Honesty the Best Policy," *Journal of Public Economics*, 2:257-263, July 1973.
- Smith, Robert: *The Occupational Safety and Health Act: Its Goals and Its Achievements*, American Enterprise Institute, Washington, D.C., 1976.
- Srinivasan, T.N.: "Tax Evasion: A Model," *Journal of Public Economics*, 2:339-346, 1973.
- Stigler, George: "The Optimum Enforcement of Laws," *Journal of Political Economy*, 78(3):526-536, May/June 1970.
- U.S. Department of Labor: *Occupational Injuries and Illnesses in the United States, 1975*, U.S. Government Printing Office, Washington, D.C., 1977.
- Viscusi, W. Kip: "The Impact of Occupational Safety and Health Regulation," *Bell Journal of Economics*, 10(1):117-140, Spring 1979.
- Zeckhauser, Richard, and Albert Nichols: "The Occupational Safety and Health Administration: An Overview Prepared for the Senate Committee on Government Operations," U.S. Senate Committee on Government Operations, August 1978.