Combining Regulation and Legal Liability for the Control of External Costs

PAUL BURROWS

University of York, York, UK
E-mail: bb6@york.ac.uk

I. Introduction

Purpose

Considering the widespread coexistence, in developed countries, of both regulatory and legal liability instruments for the control of environmental and products-related external costs, it is surprising that there have been so few economic analyses of the simultaneous use of the two instruments. More effort has gone into exploring the relative merits of regulation and legal liability as alternatives. But any attempt to establish the general theoretical superiority, in efficiency terms, of either of the instruments over the other is doomed to failure. The reason is that the advantages of each of the instruments are context specific: In some contexts one of the instruments would dominate, in others the other one would do so. Consequently, one of the instruments can be “shown” to dominate the other only by using a model with very specific assumptions.

This paper explores the simultaneous use of regulation and liability under different assumptions about how the two instruments are administered. First, in Section II it will be assumed that a regulated standard of precaution (care) and a negligence liability rule are implemented independently of each other. We will refer to this as the case of independent instruments, although it will become apparent that the consequences of the two instruments are not necessarily independent of each other, even in this case. Second, in Section III it will instead be assumed that the two instruments are characterized (as they are in practice) by a degree of evidentiary interdependence. This will be referred to as the case of interdependent instruments. Unsurprisingly, the consequences of the instruments will prove to be interdependent in this case as well. The exploration

---

1The only two formal analyses of simultaneous use are Shavell (1984b), pp. 275–278, and Kolstad et al. (1990), Section III.
2By “contexts” is meant such factors as differences in private and agency knowledge of abatement costs and differences in enforcement costs between liability and regulation regimes. See Shavell (1984a), Section II.
3White and Wittman (1983), for example, compare an idealized, perfectly internalizing liability instrument with a uniform emissions charge and a uniform regulated standard, and they claim (p. 425) that the liability rule dominates.
of the case of an evidentiary interdependence between instruments will center on two questions that have for a long time been a source of concern to legal analysts. These are whether an injurer’s compliance with a regulated standard should relieve him of liability in negligence (the compliance defense), and whether an injurer’s violation of such a standard should guarantee that he is liable for damages in negligence (the per se rule).

In considering both the independent and the interdependent instruments, attention will be focused on instrumental uncertainty, that is, the situation in which the party whose activity is being controlled, the injurer, does not know in advance the severity of one or other of the instruments to which his activity is subject.

Assumptions

The regulation and liability instruments will be evaluated in terms of their impact on the level of social cost, defined as the sum of the costs of precaution undertaken by the injurer and external damage costs suffered by the victim. The efficiency problem is to minimize social cost, and we will adopt the usual convexity assumptions required for interior private and social optima. In addition, the analysis will concentrate on the short-run consequences of the instruments by assuming zero entry-exit in the injurer’s activity. Finally, even though the analysis is not concerned with regulation and liability as competing alternatives, an attempt will be made to retain a degree of symmetry in the descriptions of the two instruments. For example, we will not assume that either regulated standards or negligence standards are necessarily optimally set.

II. Independent Regulation and Liability Instruments

Instrumental Certainty

Let $x$ be the injurer’s level of precaution and $C(x)$ and $D(x)$ be, respectively, the total cost of taking precaution and the total external damage cost. The social optimum is obtained by minimizing total social cost, $TSC$:

$$\min_x TSC(x) = C(x) + D(x),$$

which requires the satisfaction of the familiar condition:

$$C'(x^*) = -D'(x^*).$$

The optimal precaution level, $x^*$, is shown in the interconnected Figures 1 and 2 by the points A and A'.

Now let us introduce the two instruments.

---

4 The no-policy private optimum does lie at the zero-precaution corner, but we will not be concerned with this. The problems raised by nonconvexities are interesting, though they may have been exaggerated in the literature. See Burrows (1986, 1995). Cooter (1980) has argued that legal liability has the advantage of not being weakened by nonconvexities.


6 We will be assuming that violations of standards are penalized with probability 1, except where stated otherwise. To concentrate our attention on the relationship between regulation and liability the moral hazard problem of controlling the victims’ precaution levels will be ignored. We assume that compensation is moderated optimally according to any shortfall of victim precaution from its efficient level. This allows us to simplify the analysis by making injurer precaution the only determinant of external damage cost, $D(x)$, implicitly assuming victims’ precaution to be optimal.
It is not uncommon to find regulated standards modeled as inequality constraints, but this implicitly assumes that there are no violations of the standard. Instead, we will treat a standard, \( s \), as the precaution level that triggers the injurer’s obligation to pay a fine. The level of fine, \( f \), is determined by the function \( f(x) \), where \( f(x) > 0, f'(x) < 0 \) for \( x_i < s \), and \( f(x) = o \) for \( x_i \geq s \).

Consider negligence liability and a regulated standard operating simultaneously, where there is full liability assessment \( l(x) = D(x) \), and no coordination of fines and liability. The negligence due care standard may be set equal to, above, or below the regulated standard, but we first of all focus on the case of \( s_1 = \bar{x} \leq x^* \): both standards are too lenient, with the regulated standard the more lax of the two. The injurer’s liability for some form of penalty is as follows (see the three regions in Figure 2):

region 1: \( x > s_1, \ x \geq \bar{x}_1: f(x) = l(x) = o \)

region 2: \( x \geq s_1, \ x < \bar{x}_1: f(x) = o, \ l(x) > o \)

region 3: \( x < s_1, \ x < \bar{x}_1: f(x) > o, \ l(x) > o \).

The injurer’s problem is to

\[
\min_x C(x) + f(x) + l(x)
\]
by satisfying:

\[ \text{region 1: } C'(x) \geq \alpha \quad (4) \]
\[ \text{region 2: } C'(x) \leq -l'(x) \quad (5) \]
\[ \text{region 3: } C'(x) \leq -(l'(x) + f'(x)). \quad (6) \]

The injurer faces a double penalty, fine plus liability, if both standards are violated. With the configuration of standards \( s_1 < \overline{x}_1 < x^* \) the injurer chooses \( x_i = \overline{x}_1 < x^* \), thereby avoiding both a fine and negligence liability, a choice represented by points \( H, H^M \) in Figures 1 and 2. Point \( H \) is at the minimum of the total cost curve comprising:

\[
C + f + l = C + 2D \text{ for } x_i \text{ below } s_1 \\
\text{the } F \text{ to } G \text{ segment of } C + D \text{ curve for } x_i \text{ from } s_1 \text{ up to } \overline{x}_1 \\
\text{the segment of the } C \text{ curve for } x_i \text{ at and above } \overline{x}_1.
\]

The simultaneous use of the two instruments raises precaution above the regulation-only level (\( s_1 \)) toward the socially optimal \( x^* \). But the argument is symmetrical with respect to the two instruments: With \( \overline{x}_1 < s_1 < x^* \) it would be the tighter regulated standard that raised the marginal incentive to take precaution. The benefit of simultaneous use in this context is merely that it ensures that the stricter of the two instruments is in operation. Given that both \( s_1 \) and \( \overline{x}_1 \) are suboptimal, making the injurer exempt from negligence liability when he chooses \( x_i \) such that \( s_1 < x_i < \overline{x}_1 < x^* \) (regulatory compliance), or exempt from a fine when he chooses \( \overline{x}_1 < x_i < s_1 < x^* \) (compliance with negligence standard), would serve only to eliminate this benefit of simultaneous use.
As a second step in the analysis of instrumental certainty, consider simultaneous use when both standards exceed the socially optimal \( x^* \), that is, when they are too stringent. Let \( s_2 = \bar{x}_2 > x^* \). Assume the regulatory agency and the tort court set \( \bar{f}(x) = \bar{l}(x) = -D'(x) \), and that neither instrument on its own would have detrimental consequences for efficiency. In either case, the injurer chooses \( x^* \), satisfying condition (2), rather than complying with an overstringent standard. However, even in this situation two independently efficient, but uncoordinated, instruments can increase the risk of overprecaution when they are used simultaneously. The injurer who is faced with the combined penalties below \( s_2 = \bar{x}_2 \) would find that \(|\bar{f} (x^*) + \bar{l} (x^*)| > |D'(x^*)|\) and would choose a level of precaution \( x_1 > x^* \). Therefore, even in a world of instrumental certainty, the simultaneous use of uncoordinated instruments can increase the risk that the overzealous setting of standards will reduce the efficiency of the control of external costs.

Instrumental Uncertainty

The analysis of the simultaneous use of independent instruments will be completed by introducing instrumental uncertainty. Such uncertainty could relate to the levels of the standards set or to the levels of the penalties imposed for violations (fines or damages). We will concentrate on the former so that our results remain comparable with the existing literature, and, to keep the model manageable, other possible sources of uncertainty for the injurer and for the control agency will not be considered.

Under certainty the injurer knew \textit{ex ante} which of the three regions (p. 229) he would fall into for any chosen \( x_1 \), but now he must form an expectation as to which region applies. For simplicity, assume that for each of the standards the injurer, who is risk neutral, has a subjective probability distribution \( q_f(x) \) for the regulated standard and \( q_l(x) \) for the negligence standard. We also assume that the two subjective distributions have the same spread (variance) and are independent of each other: The injurer perceives that the two standards are set independently (see Section III for a reconsideration of this assumption).

For any chosen level of precaution, \( x_1 \), the injurer’s estimate of the probability of incurring a penalty of one form or the other is represented by the area to the right of \( x_1 \) under the relevant curve in Figure 3. For example, at \( x_1 \) the probability of a regulatory fine is the shaded area \( P_f(x_1) \), and the probability of liability in negligence, \( P_l(x_1) \) is one (area under the \( q_l(x) \) curve). In this example, the sum of the injurer’s expected costs at \( x_1 \) is \( C(x_1) = P_f(x_1)f(x_1) + (1) l(x_1) \), where \( f \) and \( l \) are now interpreted as the actual

---

7 It is possible that an injurer would comply with an excessively stringent negligence standard alone because compliance allows the injurer to escape all liability. We ignore this possibility and focus on the increase in inefficient compliance due to lack of instrument coordination.

8 The injurer’s uncertainty over the level of penalty could be handled in our model by replacing the \textit{actual} fine, \( f(x) \), and damage, \( l(x) \), levels by the injurer’s \textit{expected} values of \( f \) and \( l \) (for any level of precaution undertaken), conditional on the injurer being judged not to have complied with either of the standards. However, adding this second source of instrumental uncertainty for the injurer would complicate the explanation of his response while providing little extra intuition on the question of the relationship between the regulatory and legal liability instruments, which is the main concern of this paper.

The model that will be presented concentrates on the injurer’s uncertainty over the level of compliance required by any \textit{de jure} standard. A source of uncertainty that has similar consequences, but that will not be made explicit in the presentation, is the injurer’s uncertainty as to whether the agency will recognize the true level of precaution he has undertaken. If it does not, then the agency’s estimation error will lead the \textit{de facto} standard to deviate from the \textit{de jure} standard, and the \textit{de facto} standard will be more uncertain than the \textit{de jure} standard.
penalties in the event of a fine or liability being imposed. In general, the injurer’s problem is to

\[
\min_x C(x) + P_f(x) \int P(x) \, \mathrm{d}x, \tag{7}
\]

which under full damage cost-based fines and liability is equivalent to

\[
\min_x C(x) + (P_f(x) + P_i(x)) D(x), \tag{8}
\]

which requires

\[
C'(x) + (P_f(x) + P_i(x)) D'(x) + P_f(x) D(x) + P_i(x) D(x) = 0. \tag{9}
\]

How will the risk-neutral injurer’s behavior differ from that in the instrumental certainty case, where the two standards were known to him ex ante?

There has been no analysis in the literature to date of the consequences of using regulated standards and negligence liability simultaneously when there is instrumental uncertainty but the perceived uncertainty attached to one instrument is independent of the use of the other instrument. Yet, it can be shown that the impact of an uncertain instrument on the injurer’s choice of precaution level can be altered by the very presence of another instrument whether that one is certain or not. We will show this by using the case of one certain instrument, a regulated standard, and one uncertain instrument, full negligence liability. The reverse case would not radically alter the results. It will be helpful to proceed in two steps, first showing the effect of instrumental uncertainty on the precaution level induced by a single instrument operating alone, and then showing how this uncertainty effect would be altered in the presence of another (certain) instrument.
Uncertainty and the Single Instrument

When the negligence standard, $x$, is known with certainty and there is full liability but no regulation, the potentially liable injurer will exercise the lowest non-negligent level of precaution. Now consider the injurer’s response if he is made uncertain about the negligence standard, so that $q_l(x)$, with mean $\overline{\mu}$, is his subjective probability distribution for the standard. Contrary to the propositions that have been put forward by previous authors, it is not possible to make very general predictions concerning the impact of uncertainty on the level of precaution the injurer will choose. Unambiguous predictions would require strong assumptions.

The introduction of uncertainty has two effects on condition (9) at $\overline{\mu}$. First, the injurer’s subjective probability, $P_l(\overline{\mu})$, of being found liable if the $\overline{\mu}$th unit of precaution is not undertaken falls below 1, because there is now a positive probability, $(1 - P_l(\overline{\mu}))$, that the standard will turn out to be below $\overline{\mu}$. What is much less clear, however, is the relationship, if any, between the size of the probability of being found liable at $x = \overline{\mu}$ and the injurer’s degree of uncertainty concerning the negligence standard that will be set. If an increase in uncertainty is represented by an increase in the spread of the subjective probability distribution, $q_l(x)$, then we have as possibilities:

- an increase in spread that is symmetrical about a constant mean, and that leaves $P_l(\overline{\mu})$ constant at 0.5,
- an increase in spread that is biased to the left (right) of a nonconstant mean and that reduces (increases) $P_l(\overline{\mu})$, implying a negative (positive) relationship between the degree of uncertainty and the probability of being found liable at the mean-$\overline{x}$. Note that $P_l(\overline{\mu})$ is here interpreted as the probability of liability at the original mean-$\overline{x}$.

If $P_l(\overline{\mu})$ does change as uncertainty changes, then this will affect the injurer’s incentive to take precaution; for example, if an increase in uncertainty reduces $P_l(\overline{\mu})$ then ceteris paribus the positive right-hand side (RHS) of the first-order condition

$$C'(x) = -(P_l(x) D'(x) + P_l(x) D(x))$$

is reduced and the condition is not satisfied at $\overline{\mu}$. This implies an incentive to reduce the level of precaution (and $C'(x)$) as uncertainty increases, because the expected penalty from failing to take precaution at $\overline{\mu}$ falls as $P_l(\overline{\mu})$ falls. See curve (1) in Figure 4. However, it should be clear that this incentive to reduce protection is only one of the possible results of an increase in uncertainty operating through the size of $P_l(\overline{\mu})$, because this probability may rise (perhaps improbably) or stay constant (more probably). The latter is shown as curve (2).

The second effect of introducing uncertainty concerns the second of the RHS terms in condition (10). As the probability distribution spread increases, there is less concentration of probabilities at and near $\overline{\mu}$. When it is certain that $x = \overline{\mu}$, then a marginal increase in precaution above $\overline{\mu}$ has no effect on the probability of liability being imposed ($P_l(x) = 0$). Introducing a small level of uncertainty (say a small symmetrical spread of probabilities) means that there is a relatively large reduction in probability for a marginal increase in precaution, i.e., $P_l'(x)$ is relatively large. This implies a strong

---

9See Kolstad et al. (1990), pp. 893–895.
10This example is used by Kolstad et al. (1990), pp. 894–895, to derive a prediction concerning the relationship between the level of uncertainty and the amount of precaution taken. Unfortunately, the example does not seem to be consistent with the model they use, which implicitly assumes a symmetrical increase in spread.
incentive to raise precaution above $\overline{x}_m$. But it is important to appreciate that it does not necessarily imply that the injurer will need to undertake a large increase in precaution to minimize total cost, because with a small spread there is a large incentive to raise precaution only by a small amount (in the limit with vanishingly small uncertainty there is a vanishingly small amount of precaution above $\overline{x}_m$ to be expected). As the level of uncertainty increases, $P_l(x)$ falls and, ultimately, the incentive to raise precaution above $\overline{x}_m$ disappears. Assuming that for some range of uncertainty this second form of incentive does lead to protection significantly above $\overline{x}_m$ (although this is not guaranteed), then curve (3) in Figure 4 represents the second effect of uncertainty.

The overall incentive for the injurer to depart from $\overline{x}_m$ because of the uncertainty concerning the standard that will be imposed is a combination of curves (1) and (3) or (2) and (3) (ignoring a rightward bias of increasing uncertainty). Without evidence on the magnitude of the effects only the strong assumption represented by curve (2) for the effect of uncertainty via $P_l(x)$ can produce an unambiguous prediction that uncertainty will lead to overprotection, and this only for an intermediate range of uncertainty. On the other hand, an unambiguous prediction of underprotection would be made for large levels of uncertainty if there were any leftward bias in the increasing spread of the probability distribution. Of course, only if the courts set $\overline{x}_m = x^*$ can any of these

---

11Kolstad et al. (1990), pp. 894–895, predicted that large (small) uncertainty will cause underprotection (overprotection). This prediction requires, for very low levels of uncertainty, that significant overprotection will result, which will not be the case because as uncertainty approaches zero the magnitude of any overprotection approaches zero as well.
predictions of over- or underprotection be directly translated into normative statements about a failure to achieve the socially efficient level of precaution.

Some analysts may regard it as frustrating that more categorical predictions of the impact of instrumental uncertainty cannot be derived without overstretching the argument or adopting strong assumptions, although this is a common situation with efficiency theory. Nevertheless, the analysis so far may prove able to yield some further insights if we now link it with two different situations; first, the realistic one in which injurers expect that less than full liability will be imposed on negligent defendants, and, second, the simultaneous use of independent liability and regulation instruments.

Less than full liability has immediate implications for the impact of instrumental uncertainty in the single-instrument case. Relative to the situation of full liability under uncertainty, a failure of the courts to set expected liability, for any liable injurer, as high as the damage costs inflicted will reduce the overprotection or increase the underprotection. We can see this by setting total and marginal liability as \( l(x) = D(x) \) and \( l'(x) < D'(x) \), respectively, in condition (9):

\[
C'(x) = -(P(x)l'(x) + P'(x)l(x)).
\] (11)

The liability shortfall reduces the sizes of both of the positive terms on the RHS. The first of these reductions is in the marginal benefit (penalty avoided) of precaution through a reduction in liability in the event that the injurer is found liable. The second reduction is in the benefit to the injurer of reducing his risk of being found liable by raising his level of precaution. Overall, weak liability expectations increase the probability that an uncertain due care standard in negligence will lead to negligent behavior, which, if the standard were \( \overline{s}_m = x^* \), would be socially inefficient.

The Two Independent Instruments Under Uncertainty

The risk-neutral injurer’s response to an uncertain negligence standard can be altered by the presence of a regulated standard even when the subjective probability distribution attached to the negligence standard is unaffected. Assume the regulated standard, \( s \), is certain, and that \( s < \overline{s}_m \). We have seen (p. 233) that one of the effects of uncertainty in the negligence standard may be to reduce the injurer’s subjective probability (at \( \overline{s}_m \)) of being found liable. This implies the incentive, shown by curve (1) in Figure 4, to reduce precaution below \( \overline{s}_m \) increasingly as the spread of the distribution increases. In the presence of the regulated standard enforced by fines that are at least as high as damage costs, this incentive, if it exists, would be weakened in the sense that for all \( x_i < s \) it is eliminated. The nearer that \( s \) lies to \( \overline{s}_m \) the more dependent the injurer’s response to an uncertain negligence standard becomes on the second source of incentive, through \( P'(x) \) (see ps. 233–234). The support of the regulation in this situation, therefore, could reduce the risk of underprotection that could result either from uncertainty in the negligence standard biased to the left or from a shortfall from full liability. On the other hand, this support also increases the likelihood that “overprotection” could result, \( x > \overline{s}_m \), but this would be socially inefficient only if it were sufficiently large to take the level of precaution above \( x^* \).

If the regulated standard also is subject to uncertainty, the support it offers necessarily

---

12Viscusi (1991) has argued that sensational damages awards in some U.S. products liability cases disguise the fact that awards often undercompensate victims. Abel (1987) claimed that tort generally undercompensates, and Harris et al. (1984) presented a wealth of evidence to support this view for the United Kingdom.
becomes less clear-cut. We could analyze condition (9), and we would find that, with the 
\[ P_f(x)D(x) \] and \[ P_f(x)D(x) \] terms now nonzero, the outcome of the two independent 
instruments is a balance of the underprotection and overprotection incentives created 
by each instrument. But only in the particular case of a symmetrical increase in spread 
with respect to the regulated standard (in which the net incentive is to raise precaution 
above \( s \) for some range of uncertainty, so \( x \geq s \) could we be confident that the support 
provided by an uncertain regulation to a weak negligence incentive to take precaution 
would at least match that of a certain regulated standard. Where, on the other hand, an 
increase in uncertainty is leftward biased or the regulation enforcement is weak, the 
support will be less strong than with a certain, fully enforced regulatory standard.

Before turning to interdependent instruments, let us draw together the implications 
of independent instruments when they are used simultaneously. First, when the injurer 
is certain about the standards the simultaneous use of the liability and regulation 
instruments will offer the prospect of improving on the performance of the more 
efficient of the single instruments only if the instruments are weak in the sense of failing 
to enforce their own standards. As long as the standards do not exceed the socially 
efficient precaution the imposition of uncoordinated penalties involves no risk of 
overdeterrence. Second, under instrumental uncertainty with full penalties, the impact 
of the uncertainty on the injurer’s choice of precaution is unambiguous only in special 
cases. But when the penalties are less than full there is an increased probability that an 
uncertain standard will, in the single-instrument case, lead to underprecaution. When 
two independent instruments coexist, the support the one certain instrument can 
provide for another, uncertain, instrument is lost to some extent when uncertainty 
attaches also to the supporting instrument.

### III. The Simultaneous Use of Interdependent Instruments

Two questions that have long exercised lawyers have been posed for economic analysis 
by Shavell [(1984a), p. 365]:

1. Should an injurer’s *compliance* with a regulated standard relieve him of liability in 
negligence? (If the answer were yes, then it would imply in effect that \( \bar{x} \approx s \); the 
negligence standard could be no more strict than the regulated standard. This is 
the compliance defense).

2. Should an injurer’s *violation* of a regulatory standard guarantee that he is judged 
negligent and, therefore, liable for damages? (If the answer were yes, then it would 
imply \( \bar{x} \geq s \); the negligence standard could be no less strict than the regulated 
standard. This is the *per se* rule).

If the answer to both questions were yes, then it would be implied that \( \bar{x} \) should equal 
\( s \), and in practice this would usually mean that the courts would follow legislated 
standards of care in determining negligence. On the other hand, if the answer to both 
questions were no (and the same were true for the questions in reverse, for example, 
should an injurer being found non-negligent relieve him of liability to a regulatory 
fine?), then there would be no *evidentiary interdependence* between the two instruments.\(^\text{13}\)

Let us explore the two questions in terms of the consequences of evidentiary interde- 
pendence under instrumental certainty and uncertainty.

\(^{13}\)This is not strictly true because the *per se* rule could be replaced by an “evidence” rule; see note 14 below.
Assuming full expected penalties for the two instruments, it can be shown that the impacts of the compliance defense and the *per se* rule depend on the positions of the two standards, \( \tilde{x} \) and \( s \), relative to the socially efficient level of precaution, \( x^* \). This fact will help to explain some divergent views, concerning the compliance defense in particular, that are to be found in the literature.

If \( s > \tilde{x} \), the compliance defense has no effect on the impact of negligence liability because the chosen level of \( x \) exceeds \( \tilde{x} \) anyway as a result of the regulatory fine. But if \( \tilde{x} > s \), the defense strengthens the negligence rule, which would otherwise be stricter than the regulatory standard. Normative conclusions on the compliance defense naturally hinge on the \( \tilde{x}, s, x^* \) configuration prevailing. If \( \tilde{x} > s > x^* \), the compliance defense can only be beneficial; but for all other configurations it can only be either ineffective or harmful. This explains the contrast between the recommendations offered by Shavell (1984a) and Viscusi (1988). Shavell assumes \( s < \tilde{x} \leq x^* \) for at least some injurers and argues that protecting liability by denying the compliance defense provides an efficient incentive. Viscusi, on the other hand, assumes \( s \geq x^* \) and argues that liability (in particular, strict products liability) without the compliance defense will create additional, inefficient incentives for safety (precaution). Viscusi’s conclusion does not follow from his \( s \geq x^* \) assumption alone, so he must implicitly be relying also on the assumption of overcompensation, \( l(x) > D(x) \), for the inducement of overprotection.

In the United Kingdom, Parliament has refused to allow public standards to displace private rights to compensation (damages). This denial of compliance as a complete defence is also a feature of U.S. law, and it is safe, from an efficiency point of view, as long as the damages awarded are correctly set [Burrows and Ogus (1996)].

Turning to the *per se* rule, if \( s > \tilde{x} \), the rule strengthens negligence liability, but if \( \tilde{x} > s \), it has no effect.\(^{14}\) Again, normative conclusions depend on the configuration of \( \tilde{x}, s, x^* \). If \( \tilde{x} < s < x^* \) for all injurers, the *per se* rule can only be beneficial, and this is the context that favors the regulatory leadership of court decisions. On the other hand, if \( \tilde{x} < x^* < s \), the impact of the *per se* rule needs careful consideration. In these circumstances, adherence to the rule pulls the negligence standard above \( x^* \). Under the full regulatory enforcement of \( s \), with \( f'(x) = C'(x) \) at \( s \), the addition of liability under negligence at \( \tilde{x} = s \) will not alter the injurer’s choice of precaution level because he will choose \( x = s \) anyway. The overprecaution results from the overstringent regulatory standard, not from the use of the *per se* rule. If, however, the regulatory standard were not fully enforced, \( f'(x) < C'(x) \) in expected terms at \( s \), the addition of marginal liability could lead to \( f'(x) + f(x) > C'(x) \) for some \( x \), below \( s \), and then the *per se* rule’s strengthening of negligence could be held responsible for (some of) the overprecaution induced. Shavell’s rejection of the *per se* rule, therefore, must derive from the assumption either that the regulation \( s > x^* \) is underenforced, or that the marginal negligence liability is excessive, \( l'(x) > D'(x) \) for some \( x \), \( s > x \), neither of which are stated requirements for his conclusion. It seems from the analysis above that the risk of inducing overprecaution by linking the two instruments through the *per se* rule could be

---

\(^{14}\)The weaker “evidence” rule, by which the violation of a regulated standard strengthens the plaintiff’s case against the injurer but need not prove decisive, may strengthen liability if \( s > \tilde{x} \), and it will be ineffective if \( \tilde{x} > s \). The evidence rule is more flexible than the *per se* rule, more sensitive to differences in damage costs between injurers, differences in motivation, etc. In the United Kingdom the courts have preferred the evidence approach to the *per se* rule [Burrows and Ogus (1996)]; the *per se* rule has, it seems, found more favor in the United States, but is far from the universal rule there [Shavell (1984a), p. 371].
avoided if the penalties triggered by the two instruments were *coordinated* so as to ensure that \( f(x) + f'(x) \leq D'(x) \).

It is apparent that the consequences of the simultaneous use of regulation and liability can be altered by evidentiary interdependence, but whether the interdependence is efficiency enhancing depends critically on the levels of \( s \) and \( \hat{x} \) relative to each other and relative to \( x^* \). Many observers have seen both instruments as being insufficiently strong, in terms of the *expected* levels of their penalties, for various forms of external cost. In this context, say \( \hat{x} < s < x^* \), the *per se* rule is expected to be the more significant form of evidentiary interdependence, and it will add a complementary incentive to the injurer to move toward the socially efficient level of precaution.

**Uncertainty**

In their analysis of the simultaneous use of interrelated regulatory and negligence standards, Kostad et al. (1990) have offered a startling proposition. They claim that the addition of the regulated standard guarantees that any injurer will take more precaution than he would have done under negligence liability alone, *regardless* of whether the injurer’s preferred precaution level before the regulation was added would have violated or complied with the new standard.\(^\text{15}\) This implies that the joint use of the two instruments is more socially efficient than negligence liability alone if the injurer would choose an \( x_i \) below \( x^* \) in the absence of regulation, but not otherwise.\(^\text{16}\)

The categorical prediction that \( dx/ds > 0 \) for all \( x_i \) is stronger than the conclusion we were able to reach even for the case of instrumental certainty.\(^\text{17}\) The basis for the prediction is a hypothesized link between the regulated standard and the injurer’s expectation of the probability of being found liable in negligence. The expectations hypothesis used by Kolstad et al. takes a very specific form that is crucial to their strong conclusion.\(^\text{18}\) It begins from their assumptions concerning the elements of evidentiary interdependence:

1. the certain regulated standard is assumed to be fully enforced at \( s \), and the courts are assumed to operate the *per se* rule. The result is that the injurer is certain that all \( x_i < s \) will violate the negligence standard, so the “tail” of the \( \hat{x} \) subjective probability distribution below \( s \) is cut off.
2. the courts are assumed *not* to allow the compliance defense, with the result that even for \( x_i > s \) the complying injurer does not imagine that he now faces a lower (let alone a zero) risk of being found liable in negligence.\(^\text{19}\)

Having adopted these assumptions, the authors then utilize a particular formulation of the way in which the injurer incorporates his (new) knowledge of the regulated standard into his subjective probability distribution for the uncertain negligence stan-

---

\(^\text{15}\)Kolstad et al. (1990), p. 898.
\(^\text{16}\)See their Proposition 4, p. 898. Actually these are not sufficient conditions; Joint use will be more socially efficient for sure if adding \( s \) raises \( x \) and \( x \) remains below or at \( x^* \) under the two instruments.
\(^\text{17}\)If this proposition were established, then to avoid the risk of inducing overprecaution, \( x > x^* \), it may be necessary to restrict \( s \) to \( s < x^* \). See Kolstad et al. (1990), p. 899, deriving an optimal \( s^* < x^* \), a result similar to Shavell’s (1984b), p. 276.
\(^\text{18}\)See their statement on p. 897 indicating the importance of the hypothesis for their result.
\(^\text{19}\)The authors state (p. 897) that the courts do not accept compliance as a *complete defense* in negligence, but their model does not incorporate *any* such defense. This is at variance with legal practice in the United States and United Kingdom.
standard. It is supposed that the injurer will so adjust his expectations that all of the probabilities that were, before regulation, below \( s \) are transferred to \( x_i \)'s above \( s \) in such a way that all of the above-\( s \) probabilities are increased by the same multiple (illustrated in the first quadrant of Figure 5). No justification is offered for this specific formulation, but its importance for the results suggests that alternative formulations should be considered to test the generality of the conclusion. In fact, the formulation has two strange implications. First, an injurer who had chosen \( x_1 \) under negligence alone would find that when the standard is set at \( s_1 = x_1 \), so that he just complies, he would become certain that now he will be liable in negligence.\(^{20}\) Second, let \( \hat{P}(s) = 0.5 \) (say \( s \) is the mean of a normal subjective distribution). Then an injurer who had chosen a cautious precaution. However, this neat conclusion relies heavily on the strong assumptions built

under joint use will be higher than under the regulated standard alone because the

marginal increase in precaution is increased, the new Proposition would

seem to be that joint use

the two propositions concerning joint use together, the appropriate proposition would

mean of a normal subjective distribution). Then an injurer who had chosen a cautious

precaution. However, this neat conclusion relies heavily on the strong assumptions built

use, that is, as long as there is no "overshoot" as a result of the stronger incentive to take

marginal incentive to reduce the risk of liability,

under joint use will be higher than under negligence liability alone. The above analysis also suggests that precaution

than under negligence liability alone. The above analysis also suggests that precaution

under joint use will be higher than under the regulated standard alone because the

marginal incentive to reduce the risk of liability, \( P'(x_i) \), is larger in absolute terms than \( P'(x_i). \)\(^{21}\) Kolstad et al. (1990) conclude that precaution is higher under joint use

than under negligence liability alone. The above analysis also suggests that precaution

under joint use will be higher than under the regulated standard alone because the

marginal incentive to reduce the risk of liability, \( P'(x_i) \), remains positive at \( s \). Taking

the two propositions concerning joint use together, the appropriate proposition would

seem to be that joint use \( dominates \) single-instrument use as long as \( x \leq x^* \) under joint

use, that is, as long as there is no "overshoot" as a result of the stronger incentive to take

precaution. However, this neat conclusion relies heavily on the strong assumptions built

\(^{20}\)Kolstad et al. (1990), p. 898, define the new probability, \( P'(x_i) \), of the injurer being liable in negligence at \( x \), conditional on \( x > s \), as \( P'(x_i) = P(x_i) / P_i \), where \( P_i \) is the probability of negligence liability if the injurer chooses \( x = s \). Thus, if \( s = a \), then \( P'(x_i) = 1 \) and \( P'(x_i) = P'(x_i) \), but if \( s > 0 \), then \( P'(x_i) < 1 \) and \( P'(x_i) > P'(x_i) \).

\(^{21}\)The injurer's new minimization problem following the introduction of the regulated standard, \( s \), is:

\[
\min_T C(x) = C(x) + P(x)D(x)
\]

\[
= C(x) + \frac{P(x)D(x)}{P_i}
\]

The F.O.C.

\[
C(x) = \left( \frac{P(x)D'(x)}{P_i} + \frac{P(x)D(x)}{P_i} \right)
\]

is the same as equation (9), except that both of the RHS terms are inflated by being divided by \( P_i < 1 \), and that \( P_i = 0. \)
Fig. 5. Impact of a regulated standard on the probability of liability under an uncertain negligence standard; alternative hypotheses.
into the argument, especially the particular combination of legal rules assumed, as well as the particular form of expectations adjustment employed. Consider the possible injurer responses under three combinations of legal rules:

1. **Strict adherence to the per se rule, but compliance is not a complete defense.** Under the Kolstad assumptions, the new distribution, created by the addition of regulation to negligence liability, lies everywhere above the original distribution for all $x_i > s$. But with an incomplete compliance defense there will be some point, $x = d$, at which the existence of the regulated standard reduces the probability of liability, as in quadrant (2) in Figure 5. Quadrant (2) also incorporates the hypothesis that (in the range $s$ to $d$) the introduction of the standard raises the expectation of liability proportionately more for injurers whose precaution level complies with little to spare than for those further away from $s$.

Strict adherence to the *per se* rule ensures that there is no probability of liability below $s$ in both the Kolstad and the quadrant (2) case; but it is apparent that the categorical conclusion $dx/ds > 0$ does not hold for all $x_i > s$ in case (2):

- $s$ to $d$: $P_r > P_n$, $|P'_r| > |P'_n|$ both components of incentive positive
- $d$ to $e$: $P_r < P_n$, $|P'_r| > |P'_n|$ first component negative, second positive
- $e$ to $+\infty$: $P_r < P_n$, $|P'_r| < |P'_n|$ both components negative

where point $e$ is the cross-over level of $x$ at which the two distribution curves intersect.

Depending on which precaution range an injurer would have chosen to operate in under negligence liability alone, once the regulated standard is added the injurer may have an incentive to raise precaution ($o$ to $s$, of course, but also from $s$ to $d$), or to lower precaution ($e$ to $+\infty$) or may face an analytically ambiguous incentive ($d$ to $e$). Although this is less clear-cut than the Kolstad et al. (1990) categorical prediction, it does have one interesting, and arguably plausible, implication. A regulated standard that is based on good information, so that $s = x^*$, has dual consequences for injurer precaution. Not only is the (enforced) standard a minimum constraint ruling out underprecaution, it also acts as a signal that may discourage serious overprecaution. Thus, injurers who would otherwise choose an $x_i$ in the $e$ to $+\infty$ range feel more confident as result of the regulated standard and a possible compliance defense, that the negligence standard will prove not to lie in the extreme right-hand range of the original distribution. This reduces the incentive to overprotect. However, this cannot be interpreted as a case for joint use unless there is less than full enforcement of the regulated standard. Under full enforcement, the regulated standard at $s = x^*$ is optimal on its own. But joint use may dominate a regulated standard alone if the sum of the two penalties satisfies the condition $f'(x) + f'(x) = C'(x)$ at $s = x^*$. In this context the performance of the two instruments together is improved by the reduced incentive to overprotect that results from the injurer’s tendency to view the regulation as a signal to the negligence standard. But the dominance of joint use could be clear-cut only if the incentive it would create to overprotect in the $s$ to $d$ range were not significant.

---

22To avoid the proliferation of cases, we assume regulatory leadership in the sense that the certain regulatory standard is set above the best-guess level of the negligence standard, i.e., $s > x^*$. In the absence of regulation, the $x$-probability distribution is assumed to be normal on $x^*$, ranging from $x = 0$ to $s$.

23To the right of $x = d$ the area under the solid curve is less than the area under the broken curve.

24That is, the ratio of new to old probability, at any $x_i$ in the range from $s$ to $d$, is higher, the closer $x_i$ is to $s$. 

---

25BURROWS
2. No strict adherence to the per se rule, but compliance is a complete defense. This is the case in which a failure to comply with the regulated standard does not guarantee that the injurer will be found liable in negligence, but compliance does ensure that the injurer avoids liability. In quadrant (3) in Figure 5 we see that the part of the original distribution above $s$ disappears when the regulated standard is added to negligence, but below $s$ the new distribution mirrors the upper part of the new distribution in quadrant (2). The regulated standard acts as a signal and increases the probability of liability in the range from $a$ to $s$ but not below $a$, because the original probabilities above $s$ are redistributed below $s$. The probabilities of $\bar{x}$ lying close to $s$ are increased the most, and the probabilities of $\bar{x}$ near the lower tail, below $a$, are reduced.

The complete compliance defense prevents any incentive to take precaution in excess of the regulated standard ($P_s(x) = P'_s(x) = 0$ at and above $s$), and full regulatory enforcement guarantees $x \geq s$. If $s = x^*$, the regulated standard is optimal and negligence liability has no effect on efficiency. But if there is serious underenforcement of the regulation, the role of liability as an incentive to take precaution at or below $x = s$ needs to be considered. The existence of the regulated standard increases both components of the negligence incentive to take precaution in the range $a$ to $s$ (for which $P > P_s$, and $|P'_{s_1} > |P'_{s_2}$). But the result is ambiguous for the range 0 to $a$, because here $|P'_{s_2} < |P'_{s_1}$. This means that one result of the signal given to the injurer by the regulated standard is that for extremely low levels of precaution, $x_i < a$, the probability that liability will be avoided by undertaking an extra unit of precaution is small. As long as the regulated standard is effectively enforced for serious violations, $x_i < a$, the presence of negligence liability can enhance the less-than-perfect enforcement of the regulation against less serious violations, $a < x_i < s$. In this situation, therefore, again the combination of the two instruments can dominate either instrument alone if $s = x^*$.

3. Neither strict adherence to the per se rule nor a complete compliance defense. This is probably the most descriptively realistic case for the United Kingdom and United States, and analytically it comprises a combination of the cases 1 and 2 above. It should immediately be apparent that no categorical prediction such as $dx/ds > 0$ is going to be forthcoming. The change in the distribution of probabilities is shown in quadrant (4) of Figure 5, and the two components of the injurer’s marginal incentive to change the precaution level vary as follows:

0 to $a$: $P > P_s$, $\left| P'_{s_1} < \right| P'_{s_2}$ first component positive, second negative

$a$ to $d$: $P > P_s$, $\left| P'_{s_1} > \right| P'_{s_2}$ both components positive

$d$ to $e$: $P < P_s$, $\left| P'_{s_1} > \right| P'_{s_2}$ first component negative, second positive

$e$ to $+\infty$: $P < P_s$, $\left| P'_{s_1} < \right| P'_{s_2}$ both components negative

To simplify the picture, let us concentrate on the case of an optimally set, but underenforced, regulated standard $s = x^*$. There are two elements of clear-cut prediction in an otherwise ambiguous scene. First, in the vicinity of $s$, the range $a$ to $d$, the regulation gives the injurer an unambiguous incentive to raise his level of precaution. Second, for very high levels of precaution, the range $e$ to $+\infty$, the regulation provides a clear incentive to reduce precaution. As we have seen, if $s = x^*$, only the $a$ to $s$ part

---

25That $P_s > P$ for all $s_i$ between $a$ and $s$ can be deduced from the fact that the probability of $\bar{x}$ lying below $a$ has been reduced by the regulated standard: The area under solid curve is less than the area below the broken curve for all $s_i$ in the range 0 to $a$. 

---

242 Control of external costs
of the first of these two incentives is socially efficient; from $s$ to $d$ the incentive created by the regulation is to overprotect. On the other hand, the signal provided by the regulation does also reduce the incentive to indulge in extreme levels of overprecaution ($e$ to $+\infty$).

Uncertainty concerning negligence liability combined with the absence of a strict per se rule and an incomplete compliance defense introduces a strong element of ambiguity into the predicted effect of the regulation when negligence liability pre-exists. It is easy to imagine distributions for which the favorable incentives (near $s$ and above point $e$) would imply that such joint use will dominate liability alone. But contrary cases can equally well be imagined. Similarly, such joint use may or may not provide a superior incentive structure to an underenforced regulation alone.

The ambiguity could, of course, be reduced, even eliminated, by a move to a strict per se rule and to compliance as a complete defense. This would create a regulation-determined negligence incentive structure. Combining the imperfect regulatory enforcement ($f'(x) < D'(x)$ at $s = x^*$) with negligence liability imposed only for regulatory violations, would, as long as $f'(x) + E'(x) \geq D'(x)$ at $s$, lead to the full enforcement of the optimal standard. Naturally, the elimination of the ambiguity is necessarily welfare enhancing only if the regulation-dominated joint use is based on good agency information, allowing the standard to be optimally set.

IV. Conclusion

This paper has focused on the efficiency consequences of the simultaneous use of regulated standards and negligence liability for the control of external costs. Eschewing the derivation of categorical predictions through the adoption of strong assumptions, the analysis has emphasized that any conclusions based on efficiency theory alone inevitably are context specific. Bearing in mind this limitation, not to mention the fact that such considerations as transactions costs and moral hazard have not been included in the model, the following points can be made in summary.

Independent Instruments (Section II)

1. Under instrumental certainty, the simultaneous use of uncoordinated regulation and liability instruments can dominate single instruments only if the single instruments fail to fully enforce their own standards and neither of the instruments' standards exceeds the socially efficient level of precaution.
2. Under instrumental uncertainty, when the penalties are less than full, a certain instrument can support the incentive to take precaution that is provided by the other, uncertain instrument. No unambiguous predictions emerge in the case where both penalties are full; the effect then hinges on specific contexts being relevant, for example the particular form that the instrumental uncertainty takes (p. 233–234).

Interdependent Instruments (Section III)

1. Under instrumental certainty, the efficiency consequences of the simultaneous use of instruments that display evidentiary interdependence depend critically on the configuration of the negligence and regulatory standards and on the socially efficient level of precaution ($\bar{x}$, $s$, and $x^*$).
2. Under instrumental uncertainty, the particular form that the evidentiary interdependence takes has important consequences for the precaution-incentive effects of
the instruments. In what is probably the most realistic case, neither strict adherence to the *per se* rule nor a complete compliance defense, as in case 3 no categorical prediction could be made, although it could be shown how the impact of adding a regulatory standard depends on the level of precaution that the injurer would have made under uncertain negligence alone. To reduce this ambiguity of outcome would require a stricter application of the *per se* rule and of the compliance defense, thereby creating a regulation-led negligence incentive structure. This would, of course, significantly reduce the judges’ discretion in negligence cases, but as long as the penalties of the two instruments were coordinated and the regulatory standard efficiently set, the result would be unambiguously efficiency enhancing.

References


